Borrower is requisted to check the book and get the signatures on the torned pages, if any.

BASIC ELECTRONICS

ROYCE GERALD KLOEFFLER, S.M

Professor of Electrical Engineering
Head of Department
Kansas State College

With the assistance of

MAURICE WILSON HORRELL, M.S.

Supervisory Engineer
Research Laboratories
Bendix Aviation Corporation

JOHN WILEY & SONS, INC., NEW YORK
CHAPMAN & HALL, LIMITED, LONDON

PRINTED IN THE UNITED STATES OF AMERICA

PREFACE

This book has been prepared for a first course in electronics. The uthors have intended to present the subject at an intermediate level uitable for electrical engineering students who are classed as electric ower majors. It may also serve for electronic or communication najors who will continue with more comprehensive and specialized ourses in radio communication, wire communication, ultra-high fre-uencies, microwaves, and industrial electronic applications. They ave tried to present up-to-date material in a logical manner, using .S.A. and other standard symbols for diagrams and equations.

It has been the authors' desire to fill a gap in available textbooks at seems to lie between books that are elementary and those that ntain too much material to be included in a college recitation course four to six semester hours.

In preparing this book the authors found that Chapters I–VIII, clusive, and Chapter XII from Professor Kloeffler's *Industrial Elec-onics and Control* should be used again. Hence these chapters are peated here in their entirety or with minor modifications.

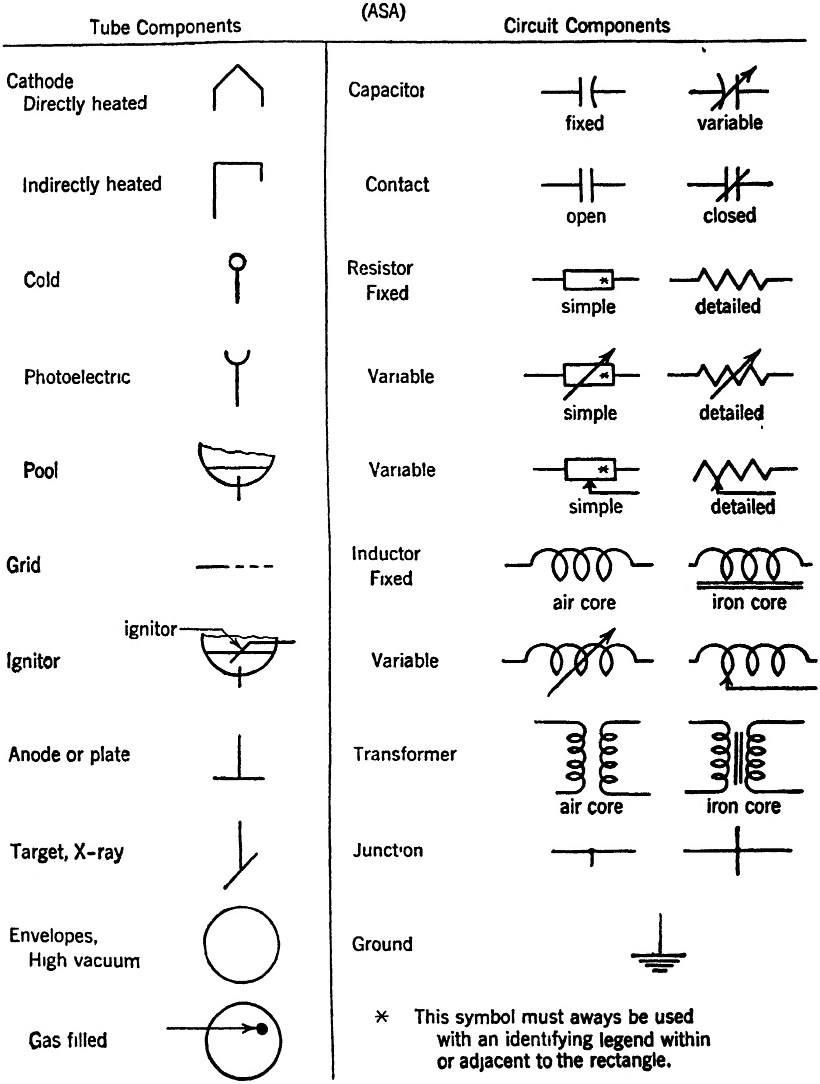
The authors acknowledge the helpful criticisms and suggestions made by Professor J. Edmond Wolfe and Robert Dennison.

Royce G. Kloeffler
Maurice W. Horrell

August, 1949

CONTENTS

ELECTRICAL GRAPHICAL SYMBOLS
(ASA)



STANDARD SYMBOLS FOR ELECTRON-TUBE CIRCUITS

Component	Grid Voltage	Plate Voltage	Grid Current	Plate Current	Load Voltage
Total value, instantaneous	e_c	e_b	i_c	i_b	e_L
Total value, instantaneous maximum	E_{cm}	E_{bm}	I_{cm}	I_{bm}	E_{Lm}
Total value, average	E_c	E_b	I_c	I_b	E_L
Quiescent or zero signal average value	E_{co}	E_{bo}	I_{co}	I_{bo}	E_{Lo}
Varying component, instantaneous	e_g	e_p	i_g	i_p	e_z
Varying component, instantaneous maximum	E_{gm}	E_{pm}	I_{gm}	I_{pm}	E_{zm}
Varying component, effective	E_g	E_p	I_g	I_p	E_z
Varying component, average	E_{go}	E_{po}	I_{go}	I_{po}	E_{zo}

Supply voltages:	
Grid (d-c)	E_{cc} or E_{cc1}
Screen grid (d-c)	E_{cc2}
Plate (d-c)	E_{bb}
Filament or heater	E_{ff}
Filament or heater terminal voltage	E_f
Filament or heater current	I_f

Chapter I

PHYSICAL CONCEPTS

Electronics is a term signifying certain developments that have centered around several discoveries and inventions made near the close of the nineteenth and the beginning of the twentieth centuries. In 1887 Hertz discovered the Hertzian waves. In 1895 Roentgen invented the X-ray tube. About 1898 Marconi demonstrated the possibilities of wireless communication. In 1902 Fleming invented the "valve" or two-electrode detector. In 1906 De Forest invented the audion or three-electrode tube. These basic discoveries gave an impetus to the work of hundreds of other scientists whose cumulative inventions and developments have produced what is known as electronics.

Electronics is that branch of science and technology which relates to the conduction of electricity through gases and in vacuo.* This definition involves the flow of electrons in vacuum tubes and the movement of electrons and ions in tubes containing gas or vapor at low pressures. The term also covers the action taking place in all circuits associated with these electron tubes. Hence, in a broader sense, electronics may be considered to include nearly all electrical phenomena.

The movement of electrons and ions in electron tubes involves certain physical phenomena not considered in the study of electrical machinery. These phenomena include (1) the removal of electrons from solids, (2) the production of ions in gases, (3) the movement of electrons and ions in space between electrodes, and (4) the control of the flow of electrons and ions by electrostatic and magnetic fields. An understanding of these processes may be aided by a review of some chemical, physical, and electrical concepts.

Atomic Structure. The electron theory of electricity and matter is a product of our twentieth-century thinking and research. At the beginning of the century Thompson suggested the electron as a part of the normal atom. In 1913 Robert Millikan published the result of his work on isolating and measuring the charge on the ion. Thus the

* Definition approved by the American Standards Association.

electron, a fundamental indivisible particle carrying a negative charge, was discovered. The counterpart or mate of the electron was named the proton and consists of a particle having a mass approximately 1840 times that of the electron and a positive charge equal in magnitude to the charge on the electron. With these two particles as the building blocks of nature, Bohr suggested the structure of the atom as shown in Fig. 1. He pictured that atom as consisting of a small dense core or nucleus about which one or more electrons revolve. This structure is analogous to our solar system; the nucleus corresponds to the sun, and the revolving electrons correspond to the earth and the planets. The nucleus of the atom contains all the protons and usually some of the electrons for the particular element involved. For the simple hydrogen atom the single proton constitutes the nucleus and the single electron the lone planet. For helium the nucleus consists of four protons and two electrons; the other two electrons revolve around the nucleus and serve as the planets of the system. The attraction of the positive nucleus for the revolving electrons is counterbalanced by the centrifugal force of their motion about the nucleus. Bohr assumed the paths or orbits of the electrons to be circles and ellipses. The electrons moving in orbits close to the nucleus have large forces acting upon them, whereas those in outer orbits are acted upon by much smaller forces. The amount of energy possessed by an electron revolving in any orbit is definite and characteristic of that orbit. To explain the property of radiation due to electrons, it was necessary to assume that an electron may have several orbits and that it is capable of jumping from one to another of these orbits under suitable excitation. The change from one orbit to another is accompanied by the absorption or radiation of energy. This radiated energy may be in the form of light, heat, or other wave energy. Because of this energy exchange some scientists have preferred to speak of the different orbits as energy levels.

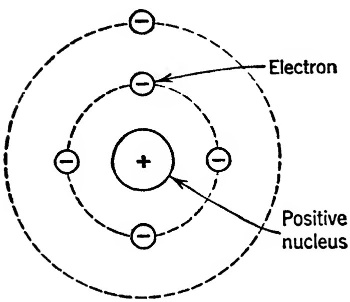


FIG. 1. Bohr's planetary structure of an atom.

The nucleus and electrons within an atom are very small and very far apart. The radius of the simplest atomic structure, the hydrogen atom, is 10^{-8} centimeter, and the radius of its orbital electron is 2×10^{-13} centimeter. The radius of the nucleus (lone proton) will be

$\frac{1}{1840}$ that of the electron. A conception of the relative magnitude of this hydrogen atomic system may be obtained by expanding the nucleus to the size of a baseball located at the geographical center of the United States (near Manhattan, Kansas). The revolving electron will pass through New York City and San Francisco and will be a sphere 300 feet in diameter—big enough to fill an average-size stadium or baseball park. Thus it is apparent that an atom is a hollow nebulous sphere—a swarm of specks occupying a small part of space. This conception is very helpful in understanding ionization in gases and many phenomena of electronics.

Two decades after Bohr's picture of atomic structure was offered, Carl Anderson discovered the positron, a positive particle having the same mass and magnitude of charge as the electron. About the same time several scientists isolated the neutron, a particle having the same mass as the proton but with a zero charge. The discovery of these two particles gave rise to new theories regarding the intimate structure of atoms and molecules. These new theories lie within the realm of chemistry and atomic physics. Fortunately for the student of electronics, the picture of atomic structure suggested by Bohr provides a very useful physical concept for the understanding of such electronic phenomena as electron emission, ionization, and light production in gases.

A molecule is usually a combination of two or more atoms.* The normal molecule contains an equal number of electrons and protons and hence the same magnitude of negative and positive charge. If an electron is removed from a molecule, the remaining unit has an unbalanced positive charge and is called a *positive ion.* If an extra electron joins a molecule, the new unit carries a negative charge and is called a *negative ion.* The subtraction and addition of electron charges in forming ions does not change the chemical nature of the molecule since the restoration or addition of an electron will bring the molecule back to its normal neutral state. It is possible for a molecule to suffer the loss or the addition of two or more electrons. In such units the particle is called a multiple charged ion of appropriate sign. A single isolated electron is often called a *negative ion,* but this terminology will not be used in this book.

Electricity. Electron theory offers a simple explanation of the phenomenon and the properties of electricity. Such explanations may

* The molecules of helium, neon, argon, krypton, xenon, and radon consist of only one atom.

be made in terms of a displacement of electrons. Thus if one or more electrons is removed from a normal (neutral) object, a *positive charge* is created on that object. Similarly, if extra electrons are added to a neutral body, a *negative charge* is created. The magnitude of the charge is measured by the deficiency or excess of electrons from the neutral state. A common unit of charge, the coulomb, consists of 6.3×10^{18} electrons. Electric charge is represented by the symbol Q.

Electric charges may be stored in capacitors (condensers). A capacitor usually consists of two parallel conducting surfaces separated by a nonconductor. A displacement of electrons from one surface to the other causes the capacitor to be charged and energy to be stored. If the two charged surfaces are later connected by a simple conductor, the displaced electrons return to their former positions, the capacitor is discharged, and the stored energy is released.

Electric charges may exist in gases or in a near vacuum as well as on metallic surfaces. In nature, clouds are often charged negatively or positively as a result of air currents and condensation of water vapor. The charges thus acquired may be great enough to result in destructive strokes of lightning. In electron tubes either negatively or positively charged ions may become concentrated in certain regions and thus constitute a charge known as *space charge.* Space charges are important considerations in the operation of electron tubes and will be covered in later discussions.

Electrons (negative charges) and protons (positive charges) attract each other. Electrons (negative charges) repel each other. Positive ions repel each other. This basic law of attraction and repulsion of electric charges is fundamental in the operation of electronic devices.

Electric potential differences are created by a displacement of electrons. If some of the electrons in a straight metal bar are moved to one end of that bar (by any means), then that end is negative and the other end is positive. A difference of potential now exists between the ends, and the magnitude of that difference depends on the density of the excess (or deficiency) of the electrons at the ends. It may also be said that electric charges exist at the ends of this rod. The magnitude of these charges depends upon the total number of electrons displaced, whereas the difference of potential depends not on the number of electrons displaced but upon the *concentration or density* of the displaced electrons. The electric charge depends on the area of the region considered, whereas the potential difference is entirely independent of the area. Potential difference is measured by the work done in carrying

a unit charge from one point to another and is independent of the path followed.

Electric conduction is the process of transferring electrons in an electric circuit. *Electric current is the coordinated movement of electrons along a conductor.* The movement may be continuous, as in direct current, or periodically changing in character, as in alternating, oscillating, or pulsating currents. The magnitude of the electric current is measured by the number of electrons that move past a point in the circuit per second. An ampere of current exists when transfer takes place at the rate of 1 coulomb (charge) or 6.3×10^{18} electrons per second. Individual electrons may move at a snail's pace in a conductor of high resistance or at a speed approaching that of light in vacuum under a very high potential. The movement of individual electrons should not be confused with the propagation of an electric wave along a conductor, which takes place at a rate approaching the speed of light. The movement of electrons called *electron current* is opposite to the conventional direction of current adopted long before the electron theory was evolved. Since electronics is a science of electron movement, the direction of current flow used in this textbook for explaining tube operation will be that of the electron current.

Electric conduction takes place in solids, in liquids, in gases, and in vacuum. In solids (metals) conduction takes place through the medium of so-called free electrons. *Free electrons* may be conceived as those electrons (1) that, while forming a part of the molecule of the conductor, lie in outer orbits and thus are not very strongly bound to the nuclei, or (2) that, at instants in their orbital movement, lie midway between the nuclei of two different molecules and hence are subject to equal but opposite attractions so that they may be readily moved by an appropriate electric field. Free electrons are moved from the sphere of one molecule subsequently to join another molecule, and thus the billions of billions of electrons in a solid conductor may move on in a kind of relay race, constituting a transfer of electrons, or electron current.

In liquids conduction is through the medium of ions. If salt (NaCl), for example, is added to distilled water, it will dissolve and disassociate into fragments or ions. The sodium atom becomes the positive ion, and the chlorine becomes the negative ion. Whether the charges exist before or only after entering solution is a question of chemistry, not electronics. If two electrodes are placed in the electrolyte and connected to a source of potential, as in Fig. 2, the ions will start to mi-

grate to the electrode of opposite polarity. When the positive sodium ion reaches the negative electrode, it will take on an electron and will become a neutral sodium atom. Likewise, the negative chlorine ion will give up one electron on reaching the positive electrode and will become a neutral chlorine atom. The result of the migration of the two ions to the electrodes has been to transfer (in effect) one electron from the negative electrode to the positive electrode, and this transfer constitutes an electron current. The formation of the neutral sodium and chlorine atom may result in subsequent chemical actions with the water present, but these reactions are foreign to the process of electric conduction in the electrolyte.

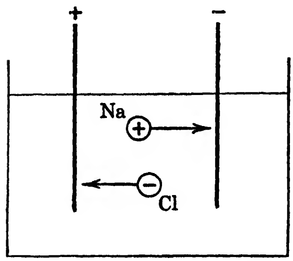


Fig. 2. Electric conduction in a solution of NaCl and water.

Control of Electron Tubes. The operation and the control of electron tubes are governed by electrostatic and magnetic fields. This statement is illustrated in the operation of the cathode-ray tube shown in Fig. 3. In this tube electrons released at the cathode C are caused to move in a stream or beam along the line COS. This motion is pro-

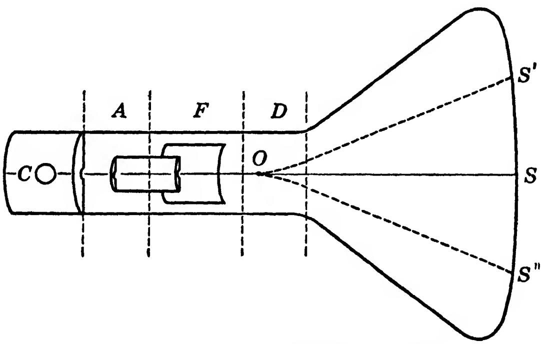


Fig. 3. Schematic cathode-ray tube.

duced by the attraction of an electric field which accelerates these electrons in the region A. Since the individual electrons repel each other because of their like negative charge, they must be held together or focused by some external force. This restraining action is applied in the region F and may be either an electrostatic or a magnetic field. In the region of D and near point O, the beam of electrons may be deflected to a new direction such as OS' or OS'', or at other angles beneath or above the plane of the page. Again, the medium of the

electron-beam deflection may be either electrostatic or magnetic fields. A more detailed explanation of the construction and operation of the cathode-ray tube will be given in Chapter IV. In order to understand its operation and that of other tubes, the theory of the movements of charged particles will be covered in a succeeding article on electron ballistics.

Mks System. The calculations in the following article on electron ballistics will use the *mks* system. In this system *length* is expressed in meters, *mass* in kilograms, and *time* in seconds. The unit of force is the *newton*—that required to give a mass of 1 kilogram an acceleration of 1 meter per second per second. This system of units is especially convenient for calculations in electrostatics since it employs the practical system of electric units—the volt for potential difference, the coulomb for charge, and the watt-second or joule for electric energy. For the magnetic circuit the unit of flux is the weber, which equals 10^8 maxwells. For this unit of flux the voltage E becomes $d\phi/dt$. Flux density β is expressed in webers per square meter.

Electron Ballistics. The theory of the movements of charged particles in electrostatic and magnetic fields is known as electron ballistics. The movements of such particles depend upon the charges and masses of the particles, the strength of the fields, and the laws of motion.

The particle of major interest in electron tubes is the electron itself. In gaseous and vapor tubes, positive ions as well as electrons are of interest. The monatomic elements of the group of inert gases and of mercury vapor are commonly used in electron tubes. In each of these, the molecule consists of one atom. Hence, either the term atom or molecule may be used in discussing the theory of the formation and movement of ions. Table 1 gives the relative electric charges and masses of several particles existing in electron tubes.

TABLE 1 *

Name	Charge	Mass
Electron	$-e$	m_0
Positron	$+e$	m_0
Proton (H ion)	$+e$	$1{,}840m_0$
Neutron	0	$1{,}840m_0$
Alpha particle (He^{++})	$+2e$	$7{,}360m_0$
Neon (ion)	$+e$	$37{,}200m_0$
Argon (ion)	$+e$	$73{,}600m_0$
Mercury (ion)	$+e$	$372{,}000m_0$

* For more precise values of m_0 and e, see R. T. Birge, "A New Table of Values of General Physical Constants," *Rev. Modern Phys.*, **13** (October 1941).

The symbol for the charge on the electron is e and for its mass m_0. Close values for the magnitude of these symbols in mks units are as follows:

$$e = 1.6 \times 10^{-19} \text{ coulomb} \tag{1}$$

$$m_0 = 9.1 \times 10^{-31} \text{ kilogram} \tag{2}$$

$$\frac{e}{m_0} = 1.76 \times 10^{11} \text{ coulombs per kilogram} \tag{3}$$

The value given here for m_0 is for the electron moving with speeds small compared with the speed of light. For higher speeds (above 15 per cent of the velocity of light), the following expression by Lorentz should be used:

$$m_v = \frac{m_0}{\sqrt{1 - \left(\frac{v}{c}\right)^2}} \tag{4}$$

where m_v = mass of the electron in motion (relativity mass).
m_0 = mass of the electron at rest.
v = speed of the electron.
c = speed of light (3×10^8) meters per second.

An **electrostatic field** is a region that exerts a force upon an electric charge. The direction of the field is that in which a positive charge is urged. An electrostatic field is produced by a change of potential E with distance.* The rate of change of potential and its sign is called the potential gradient. If a potential E_1 exists at one point and changes uniformly over a distance s to a value of E_2, the magnitude of the potential gradient is $(E_1 - E_2)/s$. It is customary to designate an electrostatic field by the term electric field intensity and the symbol $\mathcal{E}$. *Electric-field intensity* is a vector quantity that indicates the direction of the field and its magnitude as measured by the potential gradient.

$$\mathcal{E} = -\frac{dE}{ds}\dagger \tag{5}$$

* The symbol E is used because it is consistent with the standard symbols now used in electronic circuits.

† By definition, the sign of the potential gradient is negative. In the discussion and problems that follow, the negative sign will be dropped because the interest will be in magnitude only.

The strength of an electrostatic field is defined as the force exerted upon a unit charge. Hence, for a charge Q, the force will be:

$$\mathcal{E} = \frac{f}{Q} \quad \text{and} \quad f = \mathcal{E}Q \tag{6}$$

In the mks system electric-field intensity may be expressed in newtons per coulomb (equation 6) or in terms of volts per meter (equation 5), the latter being more convenient for calculations.

From the well-known laws of mechanics the following series of equations results:

$$f = ma \qquad \mathcal{E}Q = ma \qquad a = \frac{\mathcal{E}Q}{m} \tag{7}$$

where f is the force and a is the linear acceleration.

Since electric potential is measured by the work W done in moving a unit charge, it follows that for a charge Q

$$E = \frac{W}{Q} \quad \text{and} \quad W = EQ$$

Also,

Potential energy (stored) = work = EQ

Kinetic energy = $\frac{1}{2}mv^2$

Kinetic energy gained = potential energy lost

$$\tfrac{1}{2}mv^2 = EQ \tag{8}$$

$$v = \sqrt{2\frac{Q}{m}E} \tag{9}$$

Equation 9 shows that *both the speed and the kinetic energy acquired by a charged particle moving in an electric field is determined solely by the total potential* E *through which the particle has moved.* This fact gives the basis for a useful unit (electron-volt) for measuring the energies involved in particle motion in electric fields. The ***electron-volt*** is the energy acquired by an electron starting from rest and moving in a vacuum through a potential difference of 1 volt. The abbreviation for electron-volt is ev. The energy involved in electron-volts follows from equation 8.

$$\left.\begin{aligned} &\text{1 electron falling through 1 volt} = \text{1 ev (energy)} \\ &\text{1 electron falling through } E \text{ volts} = E \text{ ev (energy)} \end{aligned}\right. \tag{10}$$

Also

$$1 \text{ ev} = 1.6 \times 10^{-19} \text{ joule}$$

A uniform electric field may be produced between parallel surfaces. Let A and B represent two parallel plates separated by the distance s in Fig. 4. A potential difference E applied to the plates will set up a uniform field along the line xy and also in most of the region between the plates if the dimension of the plates is large compared to the separation s. The potential distribution along the line xy is shown in the right-hand view of Fig. 4. The potential rises uniformly (linearly) from 0 to E between the plates A and B so that the potential gradient and the electric-field intensity are constant in magnitude and equal to

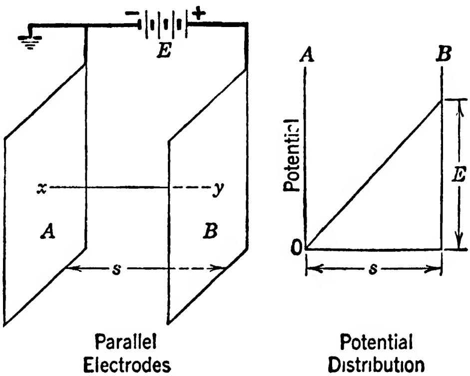


FIG. 4. Potential distribution between two planes.

E/s. An electron (negative charge) released at x will be urged toward y with a constant force equal to $\mathcal{E}e$ and accelerated uniformly until it reaches plate B at y. If E is made 500 volts and s = 2.5 centimeters, the following values of motion will result (mks system).

$$\text{Potential gradient} = \mathcal{E} = \frac{E}{s} = \frac{500}{2.5 \times 10^{-2}} = 20{,}000 \frac{\text{volts}}{\text{meter}}$$

$$\begin{aligned}\text{Acceleration (equation 7)} &= a = \mathcal{E}\frac{Q}{m} = \frac{E}{s} \times \frac{e}{m_0} = 20{,}000 \times 1.76 \times 10^{11} \\ &= 3.52 \times 10^{15} \text{ meters per second per second}\end{aligned}$$

$$\begin{aligned}\text{Speed at } y \text{ (equation 9)} &= \sqrt{2\frac{Q}{m}E} = \sqrt{2\frac{e}{m_0}E} \\ &= \sqrt{2 \times 1.76 \times 10^{11} \times 500} = \sqrt{1.76 \times 10^{14}} \\ &= 1.325 \times 10^7 \frac{\text{meters}}{\text{second}} = 4.4 \text{ per cent velocity of light}\end{aligned}$$

An electron beam (a pencil of flying electrons) is controlled and deflected by electric fields in cathode-ray tubes. To analyze such action, assume in Fig. 5 that an electron e moving with a horizontal speed of v_0 along the line aeb enters a uniform electric field (vertical) between two parallel plates at point e. Assuming the polarities indicated, it is evident that the electron will be deflected upward by the electric field and travel along the full line as shown. The speed of translation v_0 to the right will remain unchanged during the electron flight between the plates because the deflecting field is at right angles to that motion.

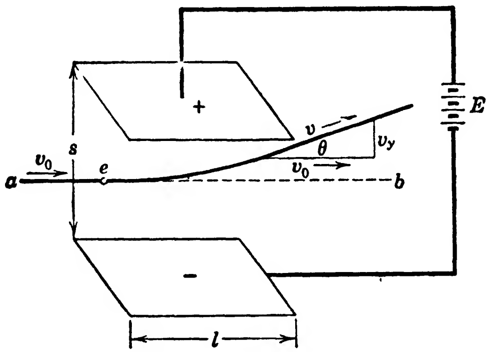


Fig. 5. Deflection of an electron in an electric field.

The electric field between the plates (assumed uniform) will give the electron an accelerated motion upward, reaching an upward speed of v_y at the time the electron has emerged from the right-hand edge of the plates. From this point the electron will move on a straight line with a resultant speed determined by the two components v_0 and v_y. The angle θ of the resultant deflection will be determined by the component velocities v_0 and v_y.

The direction and magnitude of the resulting motion can be readily calculated as follows. Assume that the initial velocity of the electron is the same as that of the previous problem (1.325×10^7 meters per second), the plates are 1 centimeter apart with 100 volts applied, and the length of the plates l is 2.5 centimeters. The transit time for the electron having an initial horizontal component speed of v_0 is derived from

$$\underset{\text{(distance)}}{l} = \underset{\text{(speed)}}{v_0} \times \underset{\text{(time)}}{t}$$

$$t = \frac{l}{v_0} = \frac{2.5 \times 10^{-2}}{1.325 \times 10^7} = 1.885 \times 10^{-9} \text{ second}$$

$$\text{Acceleration upwards} = a = \frac{E \cdot e}{s\, m_0}$$

$$= \frac{100}{1 \times 10^{-2}} \times 1.76 \times 10^{11}$$

$$= 1.76 \times 10^{15} \frac{\text{meters}}{\text{second}^2}$$

$$v_y = at = 1.76 \times 10^{15} \times 1.885 \times 10^{-9}$$

$$= 3.32 \times 10^{6} \frac{\text{meters}}{\text{second}}$$

$$\text{Angle of deflection } \theta = \tan^{-1} \frac{v_y}{v_0} = \tan^{-1} \frac{3.32 \times 10^6}{1.325 \times 10^7}$$

$$= \tan^{-1} 0.25$$

$$= 14^\circ$$

$$\text{Final speed } v = \sqrt{v_0^2 + v_y^2}$$

$$= \sqrt{1.325^2\, 10^{14} + 0.332^2 \times 10^{14}}$$

$$= 1.368 \times 10^{7} \frac{\text{meters}}{\text{second}}$$

The upward deflection of the electron while it is passing between the plates may be computed by using the formula

$$s = \tfrac{1}{2}at^2$$

$$s = \tfrac{1}{2}(1.76 \times 10^{15})(1.885 \times 10^{-9})^2$$

$$= 3.12 \times 10^{-3} \text{ meter} = 0.312 \text{ centimeter}$$

Magnetic fields are used for controlling the motion of charged particles. The action of a magnetic field upon charged particles follows from the theory of the force acting upon a current-bearing conductor placed in a magnetic field.

$$f = \beta i \sin \theta$$

where f represents force per unit length, β the flux density, i the current, and θ the angle the conductor makes with the magnetic field. A current consists of electrons or other charged particles in motion. Hence

$$i = Qnv$$

where Q is the charge on the particle, v is the speed of the particle, and n is the number of particles per unit length. Since the force on the individual particle is under consideration, n becomes unity and

$$f = \beta Q v \sin \theta$$

In Fig. 6, let it be assumed that an electron e is projected at right angles ($\theta = 90°$) into a uniform magnetic field with an initial velocity of v_0. Since the field is directed in (toward the paper), the application of any convenient rule * for the force upon a conductor will show that the electron will be deflected downward as it enters the field. With a uniform field and a constant speed v_0, the path of the electron will be given a constant angular change or acceleration and will move in the arc of a circle. From the laws of mechanics this acceleration is v^2/r, where r is the radius of curvature of the path. Thus

$$f = ma = \beta Q v = m\frac{v^2}{r}$$

and

$$r = \frac{mv}{\beta Q} \tag{11}$$

and for an electron where $Q = e$

$$r = \frac{v}{\beta} \times \frac{1}{e/m} = 5.69 \times 10^{-12}\,\frac{v}{\beta} \text{ meters} \tag{12}$$

Assume in Fig. 6 that v_0 is 2×10^9 centimeters per second and the field strength β is 10 gauss.

$$r = 5.69 \times 10^{-12}\,\frac{\overset{\text{(meters per second)}}{2 \times 10^7}}{\underset{\text{(webers per square meter)}}{10 \times 10^4 \times 10^{-8}}}$$

$$= 11.38 \times 10^{-2} \text{ meter} = 11.38 \text{ centimeters}$$

* Use left hand instead of right hand for electron current.

The theory of the magnetic field acting at right angles to the moving electron (illustrated in Fig. 6) is applied in a cathode ray tube to the deflection of a beam of electrons and is analogous to the use of the electric field as illustrated in Fig. 5.

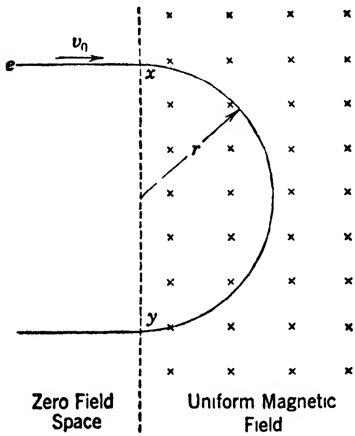


FIG. 6. Circular deflection of an electron in a uniform magnetic field.

The angle of deflection of an electron beam in a uniform magnetic field may be computed from the relations shown in Fig. 7. Here a magnetic field within the rectangle *abcd* is created by two parallel coils, one in front of and one in back of the page. An electron entering the field from the left with a velocity v_0 at point x will be deflected along the arc of a circle xy having some radius r. After emerging from the field at y, the electron will continue its velocity v_0 along a linear path. The angle of the arc xy is θ (lower angle) and by plane geometry is equal to the angle of deflection θ (at the right). Obviously, the sine of the angle θ is l/r. The length of field l is known and r can be calculated from equation 12 as in the preceding example.

If a flying electron enters a magnetic field at an angle less than a right angle, an interesting phenomenon occurs. This is illustrated in Fig. 8. An electron moving with a velocity v_0 enters a horizontal magnetic field at point e in a direction making an angle θ with the x axis. The horizontal magnetic field is produced by the solenoid $CCC'C'$. Assume for the sake of a simple concept that the magnetic field begins at the boundary CC and ends at boundary $C'C'$. The initial velocity of the electron v_0 has a horizontal component v_x and a vertical component v_y. The horizontal component v_x is parallel to and unaffected by the magnetic field. The vertical component v_y is normal to the

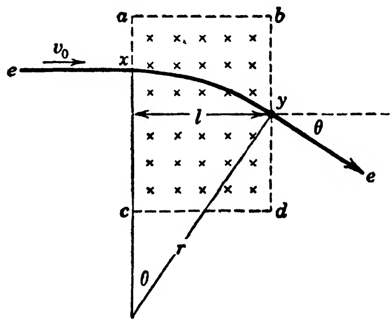


FIG. 7. Angle of deflection of an electron traversing a magnetic field.

magnetic field and hence is subject to a side thrust which will result in a circular motion as explained in the preceding discussion. The resultant motion of the electron while it is in the magnetic field will be a combination of a motion of translation to the right equal to v_x and a circular motion about a horizontal axis. Such a combination of motions follows a helical path as shown in Fig. 8. It is obvious that the effect of the magnetic field has been to change the direction of motion of v_0 into a helical motion along the direction

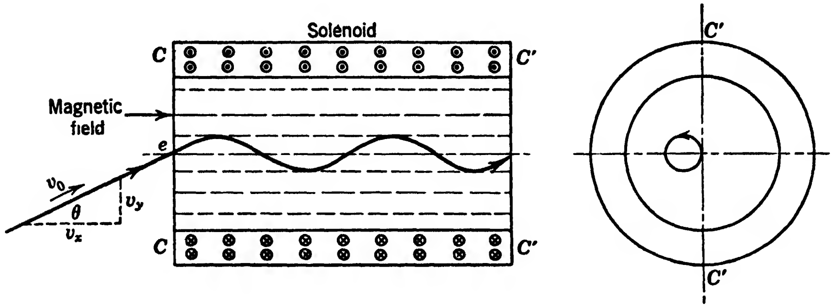


FIG. 8. Helical path of an electron after entering a longitudinal magnetic field in a solenoid (magnetic focusing).

of the field or x axis. Accordingly, a beam of electrons projected into a magnetic field at various angles other than 90 degrees will be deflected into the direction of the magnetic field. This is the principle of magnetic focusing which is applied in some cathode-ray tubes and electron microscopes. In the application of one form of magnetic focusing, the field strength is adjusted so that the electrons make one or an integral number of turns in the helical path. If all electrons enter on the axis of the magnetic field (*though at various small angles*) and make the same number of turns, they will converge on one spot. Calculations for the path of the electron in the magnetic field can be made by adapting the theory of the circular motion to the v_y component of v_0 and using the physical concept of the helix.

Concentric cylindrical electrodes are frequently used in electron tubes. A common configuration is shown in part *a* for Fig. 9. A potential E is applied to the cylinders so that the central electrode c is negative and the resulting electric field will be directed along radial lines as shown (dotted lines). In practice, electrode c may be a very fine wire giving a potential distribution along any diameter as shown in part *b* of Fig. 9. The change in potential or potential gradient close to c is very great and then becomes slight throughout most of the

radial path to the outer cylinder. If electrons are released at the wire *c*, they will fly along radial lines under the influence of the electric field to the outer cylinder.

Combined electric and magnetic fields applied to the configuration discussed in the preceding paragraph result in an interesting and use-

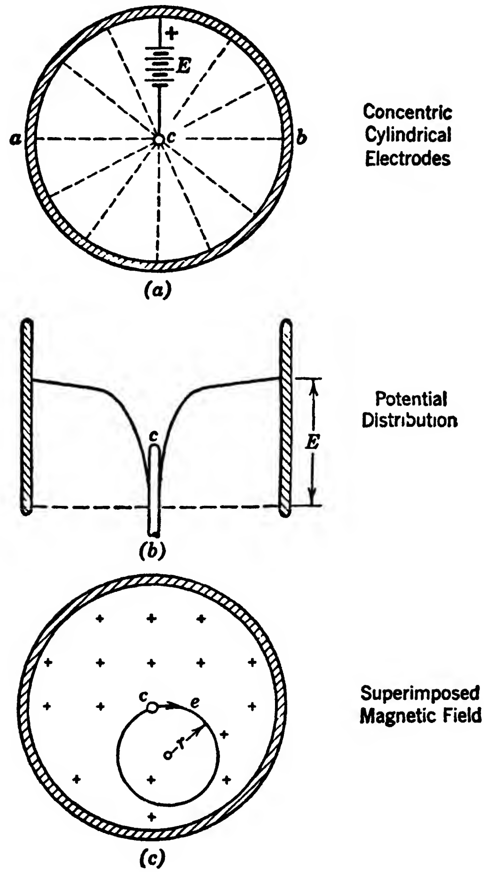


FIG. 9. Approximate path of an electron in a combined electric and magnetic field between concentric cylindrical electrodes.

ful phenomenon. If a solenoid bearing continuous current is placed around the concentric cylindrical electrodes, a field parallel to the axis of the cylinders will be created as indicated by the crosses on part *c* of Fig. 9. Now the electrons released at *c* move outward under the influence of the electric field as before, but in doing so their path is normal to the magnetic field so that they are subject to a deflection in a curved path. If the magnetic field is sufficiently strong, these electrons may

never reach the outer electrode but may be returned to electrode *c* as shown in the figure. This phenomenon has been utilized in a tube known as a magnetron and applied in an oscillator for producing very high frequencies.

In one application of the magnetron, the magnetic field is adjusted so that the electron will just graze the outer cylindrical electrode in following its curved path. The exact mathematical solution for this motion is too involved to justify treatment here.

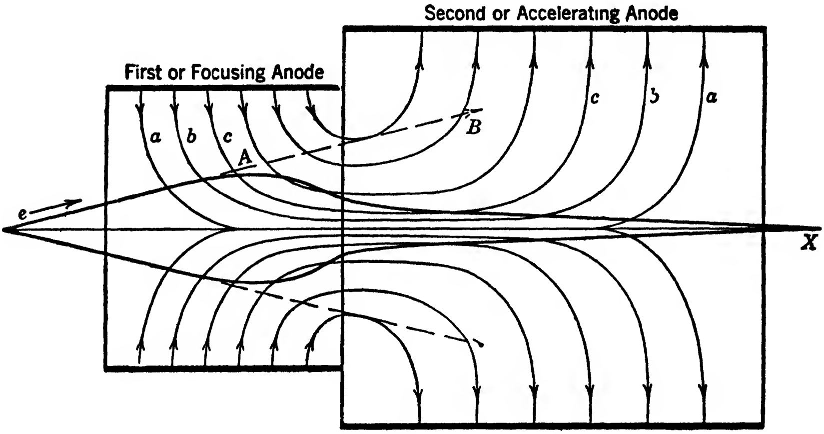


FIG. 10. Focusing of an electron beam by an electrostatic lens.

Electron optics is the science of controlling electron beams in a manner analogous to the control of light rays by optical lenses. The electron beams are directed in parallel, converging, or diverging paths through the medium of electric and magnetic fields. The focusing effect of a parallel magnetic field illustrated in Fig. 8 is an example of the methods employed. A converging or a diverging magnetic field will exercise a corresponding directive effect upon an electron beam. An electric field between adjacent concentric cylinders may be used to produce a converging or diverging effect upon a movement of electrons. This phenomenon is illustrated in Fig. 10 where the electric-field distribution shown is created in the proximity of the adjoining cylinders. Electrons moving along the general direction of the axis of the cylinders will tend to move parallel to the lines of the electric field. An electron such as *e* moving along the line *AB* in Fig. 10 will be deflected as shown, and a group or beam of diverging electrons will be directed so as to converge at a point *X* on the axis. By such methods and combinations of electric and magnetic fields, the art of electron optics is

accomplished. The complete theory of electron optics is a subject too advanced for the purpose of this book.

PROBLEMS

1. Two large parallel planes in a high vacuum separated by a distance of 1 inch carry a difference of potential of 1000 volts. An electron is freed at the negative plane (near center) with zero velocity. Calculate (a) the potential gradient between the planes, (b) the velocity of the electron when it hits the positive plane, (c) the time the electron is in transit, and (d) the energy acquired by the electron in electron volts and joules (use mks system).

2. Increase the distance between the planes of Problem 1 to 5.08 centimeters, and recalculate.

3. (a) In Problem 1 substitute a + hydrogen ion (at positive plate) for the electron and solve. Repeat for (b) a neon ion, and (c) a mercury ion.

4. In Fig. 5, an electron enters the space between the two planes with an initial velocity of 10^9 centimeters per second. If the electric field between the planes is 5 volts per millimeter, what will be the angle of flight and the resultant velocity after the electron has traveled 2 centimeters to the right?

5. In Fig. 6, an electron is hurled with a velocity of 10^9 centimeters per second into a magnetic field of 15 gauss. Calculate the nature of the path of the electron. What will be the resultant velocity 1/1,000,000 of a second after the electron enters the magnetic field?

6. An electron falling from rest in a uniform electric field acquires a velocity of 3.5×10^9 centimeters per second. What is its energy in joules? Through what potential has it fallen?

7. Substitute an argon ion with a velocity of 10^6 centimeters per second for the electron in Problem 6 and solve.

8. In Fig. 7, v_0 is 1.5×10^7 meters per second, l is 5 centimeters, and the magnetic field is 5 gauss. Calculate the angle of deflection θ.

9. In a cathode-ray tube (Figs. 3 and 7), a magnetic deflection of 20 degrees is desired where v_0 is 10^7 meters per second and l is 1.5 centimeters. Calculate β in gauss.

10. If an electrostatic field replaces the conditions of Problem 9 calculate the required potential gradient in volts per centimeter (Fig. 5).

11. Assume in Fig. 8 that v_0 is 1.3×10^9 centimeters per second, θ is 5 degrees, and β is 40 gauss (directed right to left). Calculate the path of the electron, and determine the distance CC' if the electron makes 2 revolutions.

12. A mercury ion accelerated in a nonuniform electric field hits a negative electrode with an impact of 2.4×10^{-16} joule. (*a*) What was its velocity on impact? (*b*) Through what difference of potential has it fallen?

13. Assume for Fig. 5 that the approaching electron enters the electric field at an angle of 10 degrees to the x axis (inclined upward). Set up the equations for solving θ, the final angle of deflection, and for the final velocity.

Chapter II

ELECTRON EMISSION

Electron emission is the liberation of electrons from the atomic forces in metals. Many of the physical phenomena connected with the escape of electrons were observed in the nineteenth century. In 1827 Robert Brown observed minute particles of dead matter in suspension in water under a high-powered microscope and discovered that they performed irregular wiggling motions, suggesting "life." This phenomenon came to be known as the *Brownian movement* and was explained a half century later. The Brownian movement is due to the continual bombardment of inanimate particles by the thermal agitation of the molecules in the liquid. Similar movements take place in gases and in solids. In 1883 Thomas Edison observed that the region surrounding a red-hot filament is a conductor of electricity. He placed a metallic plate in an incandescent (carbon) lamp bulb and connected it in series with a galvanometer. His observations showed that whenever the plate was connected to the positive side of the filament a current was indicated, but when it was connected to the negative side there was no deflection of the galvanometer. This phenomenon came to be known as the *Edison effect.* In 1888 Hallwachs discovered that, if he charged a zinc plate to a negative potential and then exposed it to ultraviolet light, it gradually lost its charge. In each of the experiments cited, electrons were freed from the bonds of electron affinity.

Electron Affinity. Electrons are bound within the surface of a metal by an attraction or affinity. A concept of this affinity may be attained by reference to the mechanical picture of the Bohr atom. First, one may visualize an isolated atom with its positive nucleus and the electrons revolving around it in a series of orbits at various energy levels. These electrons are bound to the nucleus by an attractive force which gives them potential energy. In their orbital motion they have kinetic energy which balances the potential energy. The electrons in the outer orbit (the valence electrons) are less closely bound and are the ones that become free electrons and participate in electric conduc-

tion and in electron liberation. Individual atoms combine into molecules held together by a displacement of the valence electrons. The molecules, in turn, are held together in the solid by a rearrangement of electrons in a form of lattice structure. Thus the electrons that may participate in any liberation are subject to an attraction or affinity due to atomic structure, molecular structure, and molecular combinations. The Bohr atom may be envisioned within a section of metal as shown in Fig. 1. An electron existing within the metal at some point such as Y will be subject to many attractions from its parent atom and from surrounding atoms. This electron will have an absolute motion consisting of its private orbital speed plus the motion of its parent atom due to thermal agitation. Obviously, the electron will occupy many positions in the course of a fraction of a second where the forces of attraction acting upon it will be balanced. While in this balanced state it is "free" in a sense, and any electric field that is present will cause it to move and to become a part of electric conduction. However, an electron in position X at the surface of the solid never reaches any balanced state but, on the contrary, is subject to all the atomic and molecular forces that bind it to the solid. This combination of forces is electron affinity. Electron affinity produces a strong potential barrier which makes it very difficult for an electron to break away from the surface of its parent solid. In order to break away, an electron must have an initial velocity or kinetic energy sufficient to overcome the barrier or electron affinity. This velocity is the resultant of absolute motion of the electron arising from the combination of electron (orbital) and thermal atomic movements. If an electron carrying a negative charge e succeeds in escaping from the surface to some position such as Z, it leaves behind a positive charge e which exerts a powerful force to bring it back to the metal. This positive charge which remains in the metal is called the ***image positive charge.***

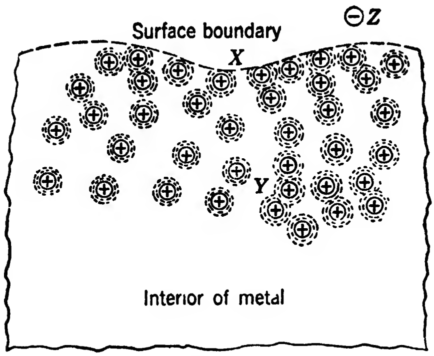


Fig. 1

The preceding discussion indicates that energy or work is required to remove electrons from solids. This energy for removing a single

electron from a solid may be designated as w. The ratio of this work w to the charge on the electron e is called the *work function* of the material. This ratio w/e is the same as the ratio of work to unit charge which, in turn, is the definition for the potential difference between two points. Thus it appears that work function or *work function equivalent* can be expressed in volts. A second way to view this subject of work function is to remember that an electron escapes from a solid through its kinetic energy. Thus

$$w = \tfrac{1}{2}mv^2$$

and the velocity v could be acquired by the action of a potential E (volts) acting on the electron with charge e. Hence

$$w = Ee = \tfrac{1}{2}mv^2$$

and work function $= w/e = E$.

To develop an understanding of work function, the student should keep in mind the following three statements. (1) Work function is a measure of the work required to overcome electron affinity plus that required to overcome the positive image charge on the surface after the electron escapes. (2) The energy of the work function is expressed in volts (electron volts), but this *does not mean* that a positive potential equal to the work function placed outside the surface will extract electrons. (3) The values of work function given in this text and in the literature refer to that *additional energy* (above that possessed at normal temperature) which is necessary to free them from the parent surface. The work function in volts for a number of materials used in electron tubes is given in Table 1. Additional discussion concerning work function will be given at the end of the chapter to explain some characteristics of the materials commonly used in thermionic emission.

TABLE 1

Material	Work Function ϕ (volts)	Material	Work Function (volts)
Platinum	5 0	Thorium on tungsten	2.63
Tungsten	4.52	Calcium	2.5
Carbon	4.5	Barium	2.0
Mercury	4.4	Sodium	1.9
Molybdenum	4.3	Calcium oxide	1.9
Tantalum	4.1	Potassium	1.55
Nickel	4.0	Strontium oxide	1.4
Copper	4.0	Cesium	1.36
Thorium	3.0	Barium oxide	1.1
Magnesium	2.7		

Electron Emission. Electrons may be liberated from metals in five different ways, as follows: (1) high-field emission, (2) secondary emission, (3) thermionic emission, (4) photoelectric emission, and (5) radioactive disintegration. Photoelectric emission consists of the removal of electrons by electromagnetic wave radiations such as light and will be discussed in Chapter XV. Radioactive materials eject electrons (beta rays) during their slow process of disintegration. This phenomenon is important in many scientific studies. In some electron tubes which contain gas and which start with cold cathodes, it is probable that radioactive materials within or near the tube and sometimes photoelectric emission provide the initial electron emission for exciting or starting the functioning of the device. The other methods for producing electron emission will be discussed in the articles that follow.

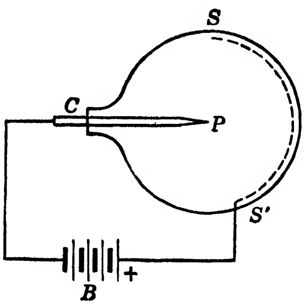


FIG. 2. An electron microscope employing high-field emission.

High-field emission is the liberation of electrons from cold metals by virtue of a very high potential gradient at the surface of the metal. At normal temperatures relatively few electrons in a metal attain a velocity at the surface sufficient to overcome the surface potential energy barrier. However, some of them do so, but upon emerging from the surface they leave behind a positive image charge which attracts them back into the surface. Accordingly, the electrons that do emerge move only an infinitesimal distance from the surface before returning. A very powerful electric field having a magnitude of over one million volts per centimeter is required to pull such electrons from the surface after this transient release. Such fields are difficult to attain, and this method finds little application in electronic devices. The student should remember the extreme difficulty of extracting electrons from cold electrodes, since this fact forms the basis of unilateral conductivity in electron tubes.

One excellent example of high-field emission has been demonstrated by J. A. Becker of the Bell Telephone Laboratories in the form of an electron microscope illustrated in Fig. 2. A spherical glass tube exhausted to a high vacuum contains a metal electrode C and a conducting fluorescent screen SS'. The electrode C is swaged to give a very fine point and is further treated by repeated dippings in an acid solution to remove any impurities and to attain a point P of nearly infinitesimal

size. A potential of 230 volts applied at B was capable of producing a potential gradient of over one million volts per centimeter at point P because of the sharpness of the point. The electrons attracted to the fluorescent screen gave a beautiful picture of the structure of the metal point P.

Another example of the use of high-field emission occurs in mercury-pool tubes where a positive space charge created by positive mercury ions produces a high field at the surface of the metallic mercury. This application will receive subsequent treatment.

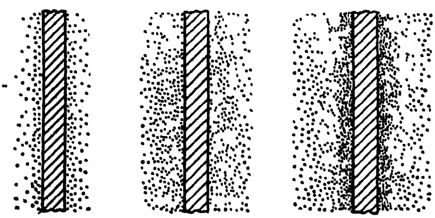
Secondary emission is the ejection or "splashing-out" of electrons from a solid due to bombardment by electrons, positive ions, or other flying particles. Usually this emission is due to electrons attracted to an electrode having a positive potential. A single primary impinging electron may eject from one to ten secondary electrons from the electrode, depending upon the work function of the material, the condition of its surface, and the velocity of the primary electron. Obviously, the kinetic energy of the primary electron is imparted to the secondary electrons and added to their normal energy in such a way that they are able to overcome the potential barrier or work function of the surface. In general, it would be expected that the energy of the impinging electron must be equal to or greater than the work function of the electrode surface where pure metals are concerned. The presence of adsorbed gas in the surface of the electrode increases the secondary emission. Here the presence of the gas molecules under the surface molecules weakens the potential energy barrier and permits the kinetic energy of the impinging electron to become more effective.

Secondary emission occurs at some electrodes in nearly all electron tubes. Sometimes its presence goes unnoted, sometimes it is a disturbing factor, and in other tubes it is employed for a useful purpose. Each case will be treated in appropriate future articles.

Thermionic emission is the liberation of electrons from a metal produced by the thermal agitation of its atoms. The phenomenon is analogous to the evaporation of liquids. At normal temperature the molecules of a liquid have some thermal agitation, but few of those at the surface "jump out" far enough to remain as vapor. With a rise in temperature the individual motions of the molecules become more violent and an increasing number do overcome the attraction of the liquid and do evaporate or "boil out." In like manner, the electrons in a metal are closely held by electron affinity, and a relatively small number have sufficient thermal energy to break away from the surface at ordinary temperatures. With a rising temperature the thermal movement of

the atoms and the kinetic energy of the electrons increase so that more and more succeed in breaking through the potential energy barrier at the surface of the metal. As the electrons break out in space, the force of attraction of the positive image charge remaining on the metal soon overcomes the initial velocity of emission, and the electrons drop back into the metal. Thus, when operating in a zero electric field, the electrons never get very far from the surface of the heated metal.

The phenomenon of thermionic emission may be conceived as illustrated in Fig. 3. Let it be assumed that a filament or cylindrical conductor of tungsten is placed in a vacuum or in a space filled with inert gas at low pressure. At a certain temperature of the filament the thermal motion of the molecules and electrons becomes great enough so that a number of electrons are thrown out as shown in part *a* of Fig. 3. A further rise in temperature of the filament will be accompanied by an increase in the kinetic energy of the electrons and an increase in the number and initial velocity of those emitted (part *b* of Fig. 3). A still higher temperature of the filament will result in a greater agitation and velocity of electrons so that a cloud or atmosphere of electrons will be formed as depicted in part *c* of Fig. 3. It should be understood that Fig. 3 represents an isolated filament and the surrounding space magnified many times. Many electrons emitted from the filament never travel over 0.01 millimeter before returning, and few go farther than 1.5 millimeters. It is obvious that the initial velocity of emission will vary over wide limits, a large number having a relatively low velocity and a few having a velocity sufficient to carry them into the outer region of the electron cloud.



(a) (b) (c)

FIG. 3. Changes of thermionic emission with temperature.

The initial velocity of emission that an electron must have in order to break away from a tungsten filament can be readily calculated from equation 9, page 8, where E represents the work function of tungsten in volts (Table 1).

$$v = \sqrt{2\frac{e}{m}E} = \sqrt{2 \times 1.76 \times 10^{11} \times 4.52}$$

$$= 1.26 \times 10^6 \text{ meters per second}$$

$$= 2{,}820{,}000 \text{ miles per hour}$$

The purity of an emitting material and the cleanliness of its surface have a considerable bearing upon the amount of thermionic emission. Wilson found that the emission of a hot platinum filament may be reduced to 1/250,000 of its normal value by first heating it in oxygen or boiling it in nitric acid. A small amount of hydrogen around a heated filament overcomes the effects of oxygen and nitric acid. The presence of a small amount of water vapor in a tube will greatly reduce the thermionic emission. Special care is used in the preparation of emitters to eliminate or neutralize the harmful effects of occluded gases and water vapors. Thermionic emission may be greatly increased by a surface layer (usually one atom thick) of thorium, barium, strontium, or calcium on a base of tungsten, nickel, and some other metals.

Equation of Thermionic Emission. Early in the twentieth century O. W. Richardson reasoned that electron emission from hot solids bore a similarity to the evaporation from liquids. Using the classic kinetic theory, he developed an equation representing electron emission and performed experiments that checked the correctness of his theory to his own satisfaction. Some contemporary scientists did not obtain equally satisfactory experimental results, which led Langmuir to refine the experiment by eliminating errors resulting from the presence of gas in the tube and impurities in the filament and thus to verify the form of Richardson's equation. Richardson suggested a second equation for emission which was formulated by S. Dushman and checked experimentally. The latter equation, which holds general acceptance today, is as follows:

$$I_s = AT^2\epsilon^{-b_0/T} \tag{1}$$

where I_s = the emission current in amperes per square centimeter of emitting surface (saturation current density).

T = the absolute temperature (degrees C + 273) = Kelvin.

ϵ = the natural base of logarithms.

A = a constant.

$$b_0 = \frac{\phi}{k} = \frac{\text{Work function}}{\text{Boltzmann's constant } (0.863 \times 10^{-4} \text{ volt per degree K})}.$$

The various factors that control electron emission from hot bodies have made it difficult to correlate the values of constants derived from scientific deductions with those determined experimentally. Accordingly, the constants are determined by experiment and calculations, and the equation is treated as empirical.

For pure metals A is a universal constant, but for coated metal cathodes both A and b_0 vary because of the degree of coverage of coated material, temperature, and other factors. Experimental values of the constants for some cathode materials are given in Table 2.

TABLE 2

	A	b_0
Tungsten	60.2	52,400
Thoriated tungsten	3.0	30,500
Barium oxide	0.01	12,000

The curve representing the trend of Richardson's equation is given in Fig. 4. It should be noted that this curve and equation 1 give the emission current (saturation current density) from a hot body or cathode and do *not* represent the current that may pass to a plate or anode in an electron tube.

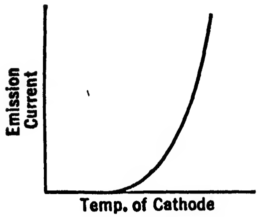


FIG. 4. Curve of thermionic emission.

Construction of Cathodes. The electrode from which electrons pass into the vacuum or gas in an electron tube is called the *cathode*. Cathodes may be classified as hot or cold. Hot cathodes may be either directly heated or indirectly heated. The directly heated type uses a filament construction as illustrated in parts *a, b,* and *c* of Fig. 5. The filament of parts *a* and *b* consist of a flat ribbon or round wire which is heated by current passing throughout. Part *c* shows a wire filament wound into a fine helix. The helical construc-

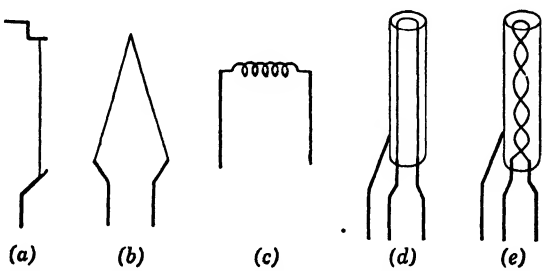


FIG. 5. Construction of typical cathodes for vacuum tubes.

tion is also used for other configurations such as in parts *a* and *b*. The material for the emitter of the filamentary cathodes of *a, b,* and *c*, Fig. 5, may be pure tungsten, thoriated tungsten, or an oxide-coated alloy of nickel. The cathode of the indirectly heated type is illus-

trated in parts *d* and *e* of Fig. 5. It consists of an oxide-coated sleeve or tube which constitutes the cathode proper plus a loop or twisted tungsten heating unit placed inside the sleeve. The tungsten heater is insulated from the sleeve by a coating of aluminum oxide. The cathode heater tube is usually made of nickel or some nickel alloy such as Konal—an alloy of nickel, cobalt, iron, and titanium.

Tungsten Cathode. The desirable properties of a material for a thermionic emitter of electrons are a high melting point, a low work function, and a long life. Tungsten has been used as an emitter for many years. It melts at 3600 degrees K and is generally operated at temperatures from 2450 degrees K to 2600 degrees K. At these temperatures it can be operated for several thousand hours to furnish an excellent supply of electrons. Tungsten has a relatively large work function so that its efficiency as measured in amperes of emission per watt of heating power is rather low. Pure tungsten was used in early electron tubes of all kinds and is still used in tubes that have a high plate voltage (above 10,000) and wherever severe positive ion bombardment of the cathode is likely to occur. The characteristics of pure tungsten as an emitter are shown graphically in Figs. 6 and 7.

Other metals that have been used sometimes for cathodes are tantulum, columbium, platinum, and molybdenum. Tungsten is preferred because of its mechanical strength and because it can be obtained more readily.

Thoriated-Tungsten Cathode. The thoriated-tungsten cathode was developed by Langmuir and his co-workers. This cathode is made of pure tungsten impregnated with approximately one per cent of thorium oxide (thoria). At its normal operating temperature of 2000 degrees K, the uncarbonized thoriated-tungsten cathode has an emission per unit of surface more than 10,000 times that of a pure tungsten filament. And its emission (at 2000 degrees K) is more than 90 times that of tungsten at the temperature of 2400 degrees K.

The thoriated-tungsten cathode is prepared for service by being mounted in a tube and heated to 2700 or 2800 degrees K for a few minutes. It is then operated at the normal operating temperature of 1900 to 2000 degrees K. The first heating at the higher temperature reduces part of the thorium oxide to pure thorium. The thorium atoms thus formed diffuse through the body of the cathode and slowly come to the surface where they form a skin or layer one molecule thick. Thus the emission of electrons comes from the thin layer of thorium atoms which have a low work function and hence give a copious emission. At the normal operating temperature, the thorium atoms are

evaporated slowly from the cathode but others from the inside diffuse to the surface to take their place. If the thoriated cathode is raised to a higher temperature, the electron emission increases but the rate of evaporation of thorium also increases and, at a certain critical temperature of about 2200 degrees K, all the thorium atoms leave the surface. Since the diffusion from the inside cannot be rapid enough to replace

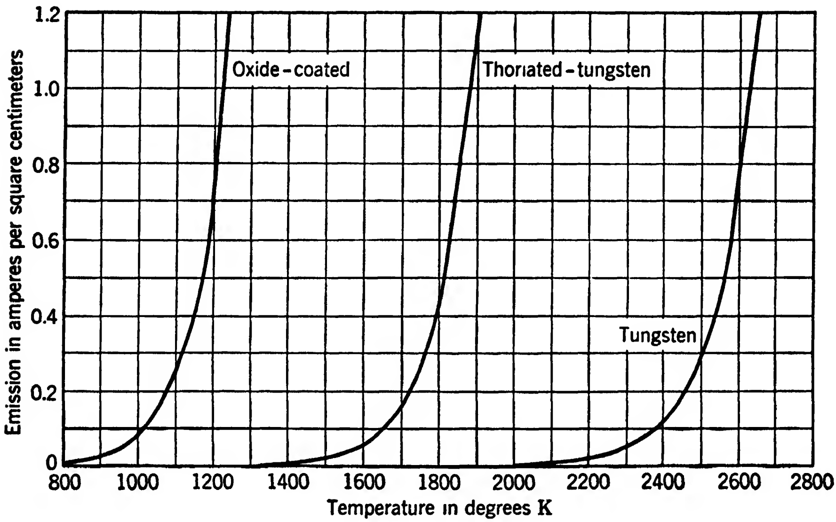


FIG. 6. Emission of typical hot-cathode emitters.

them, the emission drops to the value for pure tungsten. Thus the temperature of operation of this emitter is important in order to maintain the proper equilibrium between the thorium supplied from the interior and the amount distilled from the surface.

After thousands of hours of service, all the thorium atoms within the cathode may be used for replacements and then the emission of the cathode will fall off. In general, the filament or cathode may be rejuvenated by heating it for a short time to 2800 degrees K and then by restoring the temperature to 2000 degrees K. The first step reduces more thorium oxide to thorium and probably drives from the filament certain impurities, such as gas atoms, which have become occluded in it during the period of operation. This step is sometimes called deactivation since it stops the emission of the thorium for the time being. The second step is known as activation. Under activation the normal diffusion of thorium restores the thorium atomic layer on the surface and within a few minutes the emission returns to its full value. The

quantity of thorium oxide in the cathode is usually sufficient to repeat the deactivation and activation process several times.

The thoriated-tungsten cathode is sensitive to bombardment by positive gas ions which knock off the thorium surface atoms and reduce the

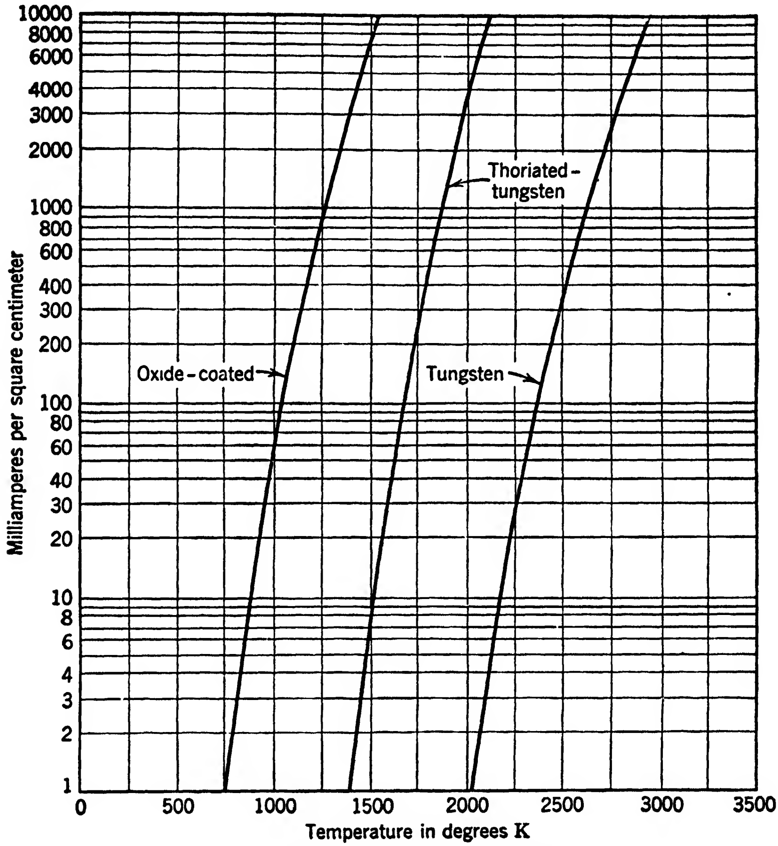


FIG. 7. Thermionic emission of cathodes.

emission. This sensitivity has been reduced by a carbonizing treatment which is given to all thoriated-tungsten cathodes manufactured today. Two methods have been employed for carbonizing. In one method the cathode is placed in a vacuum for cleaning; acetylene gas under low pressure is then introduced and the cathode is "flashed" at a temperature of 2200 to 2300 degrees K. In this process the surface of the cathode is reduced to tungsten carbide, W_2C, which has a higher resistance than pure tungsten. During "flashing," the cathode is held

at a constant impressed voltage until the current decreases by 7 to 10 per cent of its original value. This decrease in current occurs when 20 to 30 per cent of the cross section of the cathode has been converted to W_2C. Further carbonization is undesirable because W_2C is brittle and weakens the cathode to mechanical shock. Carbonization at a temperature lower than that suggested results in the formation of WC, which is unsatisfactory for the desired purpose. Carbonization can be effected by cleaning the cathodes in hydrogen vapor and then flashing at atmospheric pressure in nitrogen containing a low percentage of hydrocarbon vapor such as hexane or benzene, with the same control as suggested previously. While 4000 volts has often been suggested as an upper limit for plate voltage for thoriated tungsten, successful operation up to 10,000 and even 15,000 volts has been secured in carefully evacuated transmitting tubes. Normally, the carbonized cathode does not require any additional treatment for activation; when treatment is needed, it is the same as suggested earlier.

The emission characteristic of the thoriated-tungsten cathode is illustrated in the curves of Figs. 6 and 7. This emitting material (uncarbonized) was widely used in the construction of tubes for radio receiving sets during the period 1925–1930. The general adoption of the heater type of cathode after that period has confined the application of the thoriated-tungsten cathode to power amplifier tubes operating in the range of 750 to 5000 volts.

Oxide-Coated Cathode. In 1904 Wehnelt discovered that a small amount of calcium oxide on platinum greatly increased the electron emission. This discovery led to many experiments using various alkaline earth oxides placed on metallic cores and to the development of excellent cathodes. Barium and strontium oxide were found to be excellent materials; a mixture of approximately 50 per cent barium oxide and 50 per cent strontium oxide provides a coating with satisfactory mechanical properties and a copious emission.

The oxide-coated cathode is analogous to the thoriated-tungsten cathode in a number of ways. First, the addition of the oxide increases the emission as does the thorium coating. Second, the oxide-coated cathode has a high emitting efficiency like the thoriated-tungsten cathode. Third, any overheating of the oxide-coated filament above a critical temperature greatly reduces the emission. Fourth, the oxide-coated filament must be activated before it is ready for service.

The base or core for the oxide-coated cathode is generally made of nickel because of its good physical properties and low cost. Several alloys have been developed for the cathode base, among which Konal

offers a speedy chemical reduction of the oxide, giving a supply of the free alkaline earth metal for activation. The alkaline earth oxides are not stable in ordinary air; hence it is necessary to coat the core with their carbonates, nitrates, or hydroxides and subsequently to reduce these to oxides. The barium and strontium carbonates or nitrates may be applied to the emitter base by dipping the filament in a water suspension of the salts or by spraying the base with a suspension in amyl acetate with a little nitrocellulose for a binder. For tungsten filaments, hydroxide coatings may be applied by dipping the filaments in a bath of molten hydroxide. After the coating is applied, emitters are dried in air by heating.

The prepared filament is now mounted in the tube in which it is to be used. The tube is evacuated and kept on the pumps while the cathode is heated to a temperature of about 1400 degrees K to reduce the coating to oxide. Unless some powerful reducing agent has been present, the filament will still be "unactivated." This means that its electron emission at the working temperature of 800 to 1000 degrees K is still very low. Activation is now brought about by (1) prolonged heating, (2) applying a potential of several hundred volts to the anode, or (3) both. When the electron emission becomes about 300 milliamperes per square centimeter for a temperature of 1100 degrees K, there will be no further increase and the filament is ready for use. Activation means that free barium or strontium has been liberated from the oxide by chemical action, by electrolysis, by positive ion bombardment, or by a combination of these processes. The free metal atoms are distributed throughout the body of the oxide and on its surface in such a way as to induce a copious electron emission from the outer surface. It is generally agreed that the emission comes from the metallic atoms and not from the oxide or the base of the cathode.

After several thousand hours of normal use the emission from the oxide-coated cathode declines rather rapidly. When this point is reached the tube should be discarded since little can be done to restore its emitting property. Tubes using oxide-coated cathodes are frequently guaranteed for a life of 8000 hours and commonly operate for 20,000 hours and longer. The temperature emission characteristics of oxide-coated cathodes are shown in Figs. 6 and 7.

Oxide-coated cathodes are made in both the filament and heater types illustrated in Fig. 5. The operating temperatures for these cathodes vary from about 800 to 1100 degrees, depending upon the emission required, the expected life, and the application.

Effect of Gas upon Emission. The presence of gases has a harmful effect upon the emission and operation of *vacuum* tubes. It has been pointed out that oxygen and water vapor greatly reduce the emission from some materials and that the bombardment by positive ions may destroy the emitting surface layer of coated cathodes. Thus it is imperative that harmful gases be removed and, in the case of vacuum tubes, that all gases be removed as far as feasible. The gas that is present in a tube comes from two sources. First, there is the gas that occupies the open space in the tube, and, second, the gas that is occluded or adsorbed in the surface of the metal, glass, and other materials inside the tube. The occluded or adsorbed gas can be removed by heating the parts where it resides. This heating may be accomplished by passing current through the material as in emitters, by placing the tube in a high-frequency induction field, or by electron bombardment of electrodes. Oxygen may be removed by using hydrogen gas or by exhausting the tube and then refilling the space with an inert gas. The final evacuation is accomplished in two ways. First, a vacuum pump removes a large part of the gas, and then a material known as a "getter" completes the operation. The getter is a chemical substance, such as barium, magnesium, aluminum, and tantulum, that has the property of combining with gases when they are vaporized. In glass tubes a small amount of getter is mounted in a position where it will be heated and vaporized by the induction field. The effectiveness of the getter results from its chemical combination with the gases during flashing and from subsequent absorption of the gases by the residue of the getter which is deposited on the walls of the tube. The methods for the removal of gas are usually combined in various ways in the manufacture of vacuum tubes.

Comparison of Emitter Materials. The characteristics of emitters vary so widely with the temperature and other factors that it is difficult to make comparisons. An approximate comparison of emitting

TABLE 3

Cathode Material	Temperature (Degrees K)	Amperes per Square Centimeter	Milliamperes per Watt	Watts Required per Ampere
Tungsten	2450–2600	0.2–0.65	3–8	333–125
Thoriated-tungsten	1900	1.15	100	10
Oxide-coated	800–1100	0.01–0.22	50–300	20–3.3

efficiency is given in Table 3. A slight increase in operating temperature will increase the efficiency of emission at the expense of the ex-

pected life. At the optimum condition of operation a thoriated-tungsten emitter will give approximately 1 ampere of useful emitted current per square centimeter of surface. Under similar conditions the oxide-coated emitter may give only 0.1 ampere or less of useful current. It might be concluded from this comparison that the thoriated tungsten was more desirable; this would be true if the volume of emitted current per unit area were the important factor. On the other hand, an oxide-coated emitter will require approximately one-tenth as many watts energy per square centimeter for operation as a tungsten cathode. This comparison points to the use of oxide coating for low exciting watts and small emission currents. It should be remembered that only tungsten will withstand positive ion bombardment where very high voltages are involved.

The thoriated-tungsten filament is suitable for operation at fairly high voltages and is capable of being reactivated after the emission has been lost owing to temporary overload. The oxide-coated cathode has very high emitting efficiency, but it has a tendency to contaminate adjacent electrodes with small quantities of active emitting material so that emission may take place from these electrodes at relatively low temperatures.

There is no best among the various emitting materials and each has its appropriate place in the field of application, depending on the requirements of operation and the economy of design.

Applications of Emitter Materials. The oxide-coated cathode must always be used for the indirectly heated type because the other cathodes require too high a temperature for satisfactory operation. This application covers nearly all the millions of tubes in use in alternating-current receiving sets. The oxide-coated filament type is best suited wherever quick "heat up" is necessary, and it is essential where a low heating energy drain is necessary because of long hours of operation. This application covers the thousands of tubes in service as telephone repeaters on the toll lines in America.

Thoriated-tungsten and tungsten cathodes find their application in the power tubes used in radio, telephone, and carrier-current transmitters.

In general, it may be said that (1) oxide-coated cathodes are used for tubes up to 100 watts with plate ratings up to 750 volts,* (2) thoriated-tungsten cathodes are used for tube capacities from 100 to

* High plate-voltage ratings up to 20,000 volts have been used in gaseous tubes and large vacuum tubes having sufficient cathode-anode spacing.

1000 watts with plate voltages up to 4000 and 5000 volts, and (3) tungsten is used for cathodes in tubes with ratings from 1 kilowatt up and plate potentials of 5000 volts and higher.

Theories of Work Function. The emission characteristics of cathode materials are more readily understood when viewed in the light of theories of metals suggested in the 1930's.* Under earlier concepts the electrons in metals at ordinary temperatures were assumed to have low energy values. Accordingly, the work function ϕ was assumed to be the total work necessary to remove an electron from a metal. The newer theories advanced by Sommerfeld, Fermi, and others have indicated that the electrons have much higher energies than previously considered. The maximum energy of the electrons in a metal at low temperatures is found to be of the order of 10 volts. This new value has been confirmed by experiments of Davisson and Germer for the reflection of electrons from crystal surfaces. The new findings suggest a simple concept for emission, as shown in Fig. 8. Here the vertical distances indicate the energies of electrons in electron-volts. Within the metal the electrons may have energies varying from zero to μ. In order to escape into the vacuum outside the surface boundary, a total energy equal to W is required. The difference in energy $W - \mu = \phi$ is ordinarily termed the work function of the material and represents the values given in Table 1.

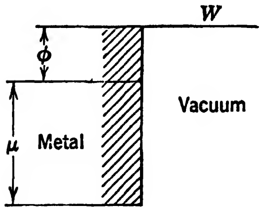


FIG. 8

The concept represented in Fig. 8 for the surface potential boundary is an ideal one, but it fails to account for the effect of the image force which is exerted as soon as the electron emerges. The action of the image charge is shown in Fig. 9 for a small magnified area on the emitting surface. Whenever an electron of charge e emerges from the surface by a distance x, it is subject to a force due to the image charge $+e$, in accordance with Coulomb's law.

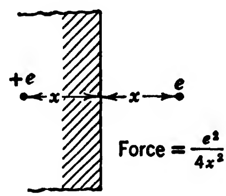


FIG. 9

$$\text{Force} = \frac{e^2}{4x^2}$$

* Much of the material in this article has been taken with permission from *Notes on Industrial Electronics* by I. E. Mouromtseff and D. E. Marshall of Westinghouse Electric Corporation.

To move the electron away from the surface against this force requires a varying amount of energy. Since the surface is not smooth, the distance x is hard to define for small values (a few atomic diameters). Hence it is permissible to assume that for small values of x the force will vary linearly, and that for larger values it will vary in accordance with Coulomb's law. This assumption will give a more realistic concept of the energy required, as shown in Fig. 10. The ordinates of the curved boundary line on the right indicate the total energy (ev) that the electron must acquire to move to the given distance from the surface when in free space (zero electric field). Here W again represents the total energy required for emission.

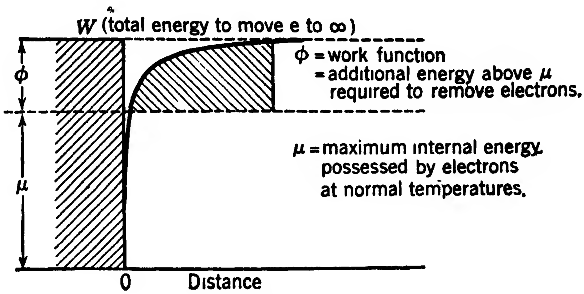


FIG. 10

The emission from hot cathodes does not follow Richardson's equation when an electric field is applied in the surrounding region. Schottky showed that this phenomenon was due to a lowering of the work function or the surface potential barrier by the external electric field. The phenomenon is known as the Schottky effect. Schottky derived the following equation:

$$I = I_0{}^{e^{3/2}E^{1/2}/KT} \tag{2} *$$

where I_0 = current from Richardson's equation with zero field. This equation agrees with experiment to within a few per cent for fields up to 10^6 volts per centimeter for smooth pure metal surfaces. Under normal operating conditions where the applied field is of the order of 2000 volts per centimeter, the increase in emission due to this field is about 10 per cent at 2000 degrees K. A graphical concept of the Schottky effect is given in Fig. 11 which follows from the discussion of Fig. 10. The ordinates of the line labeled "applied field" show the energy given to the electron by the applied field as the electron moves away from the surface. Obviously, this energy must be subtracted

* E is electric field at cathode surface.

(dotted line above) from the energy required for escape in a zero field, giving a net energy shown by the full line. The maximum height of the escape curve is W', leaving a difference $W' - \mu = \phi'$, which is the extra energy or work function for the conditions assumed.

The preceding theory showing the effect of an external applied field can be employed to explain the phenomenon of high-field emission. In

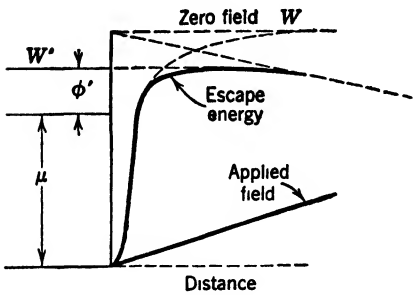


Fig. 11

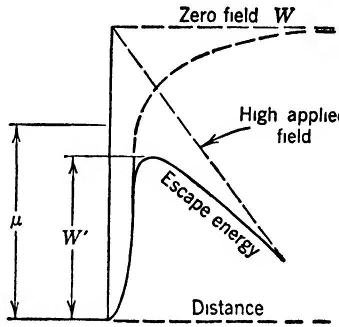


Fig. 12

Fig. 12, a very high field is shown by the dotted line. When the force of this high field is acting upon an escaping electron, the energy that must be possessed by the electron is indicated by the full-line energy curve. Since the maximum height of this energy curve is W', it is apparent that those electrons within the surface of the metal that have

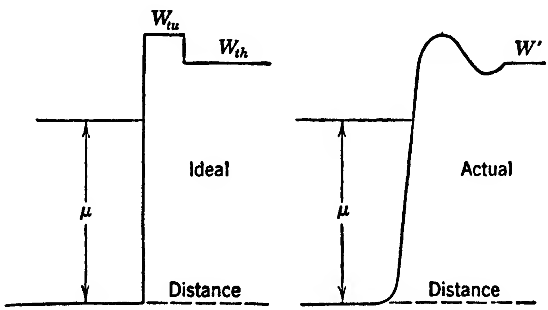


Fig. 13

energies between W' and μ may readily escape even at normal or low temperatures. This is high-field emission. Such emission usually gives large currents which are likely to be destructive to the cathode surface.

The preceding concepts of work function at surfaces can be applied to other than pure metal cathodes. For the thoriated-tungsten cathode

the potential barrier at the surface is pictured in Fig. 13. The ideal arrangement on the left shows the higher work function of the tungsten W_{tu} and the lower one of thorium W_{th}. The ideal picture is modified by the image force at the surface, giving the more probable potential barrier relation shown on the right. Apparently the electrons possess sufficient energy to carry them through the hump with enough remaining energy to escape after emerging into the lower work function region. Another somewhat obscure factor known as polarization

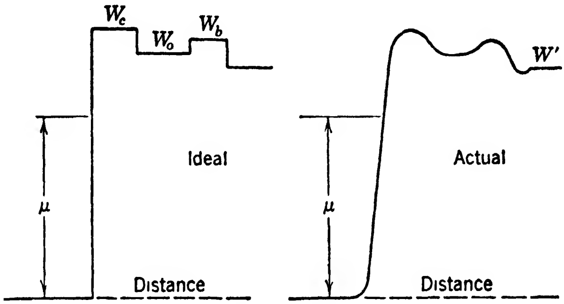


FIG. 14

enters into the picture. Polarization means that the thorium atoms on the surface of the tungsten reorient themselves in such a manner as to reduce the value of the work function W' below that of thorium on thorium. Note also that an external field will become more effective in reducing the final value of W' after electrons pass over the hump. This explains the fact that the emission from both the thoriated-tungsten and oxide-coated cathodes is much more sensitive to the effect of external fields. Similar diagrams for the required emission energy for oxide-coated cathodes are given in Fig. 14. Here the emitted electrons must pass through two potential barrier humps in order to escape. Electrons in the process of emission must overcome the work function of the core W_c, the oxide W_o, and the surface barium atoms W_b before liberation is achieved.

PROBLEMS

1. What must be the initial velocity of an electron in centimeters per second in order to be emitted from (a) thorium? (b) thoriated tungsten? (c) barium? (d) barium oxide?

2. Assume that a certain tungsten cathode emits 1 milliampere of current at 2100 degrees K. Calculate and plot the curve of its emission through points 1800, 2000, 2200, 2400, 2600, and 2800 degrees K, using Richardson's equation.

3. Repeat Problem 2 for a thoriated-tungsten filament emitting 1 milliampere at 1500 degrees K for a range of 1400 to 2200 degrees K.

4. Repeat Problem 2 for a barium oxide filament emitting 1 milliampere at 800 degrees K for a range of 800 to 1200 degrees K.

5. A certain tungsten cathode is operated by an applied potential of 5.0 volts and 1.0 ampere and supplies a saturation emission current of 50 milliamperes. What per cent of the input is transformed into effective energy of emission?

6. A thoriated-tungsten cathode emits 1.15 amperes per square centimeter at 1900 degrees K. What will be its emission at 2300 degrees K?

7. Using Table 1 and equation 1, calculate the curve of emission for tantalum for temperatures between 2000 and 2800 degrees K.

8. A tungsten filament has a diameter of 0.025 inch and operates at 2500 degrees K. What must be the length of the filament to give a saturation current of 0.5 ampere?

9. A barium oxide-coated cathode consists of a cylinder 0.1 inch in diameter and 1.0 inch long. What will be the saturation current when the cathode operates at 1000 degrees K?

10. Calculate the ratios of the saturation emission currents from tungsten, thoriated-tungsten, and barium oxide-coated cathodes for normal operating temperatures of 2400, 1900, and 1000 degrees K.

REFERENCES

Richardson, O. W., *Emission of Electricity from Hot Bodies*, Second Edition, Longmans, Green & Co., 1921.

Langmuir, I., "The Electron Emission from Thoriated-Tungsten Filaments," *Phys. Rev.*, **22**, 357 (1923).

Dushman, S., "Thermionic Emission," *Rev. Modern Phys.*, **2**, 381 (1930).

Chapter III

VACUUM DIODES

Unilateral Conductivity. Edison's early experiment in which he observed the "Edison effect" showed that electrons would pass from a hot filament or cathode to a cold plate whenever the plate was at a positive potential with respect to the filament, but that no electrons would pass in either direction if the plate was at a lower potential (negative) with respect to the filament. This property of a device which permits electrons to flow in one direction only is called *unilateral conductivity*. The unilateral conductivity of a tube having two electrodes, one hot and the other cold, was utilized by Fleming for the detection of high-frequency radio waves. This device, known as the Fleming valve, was patented by him in 1905. The Fleming valve was important in the early application of radio telegraphy, and it was one of the leading discoveries in the history of electronics.*

Contact Electromotive Force. Volta discovered that, when two different metals were placed in contact and then separated, they acquired electric charges. He also discovered that, when two different metals were placed in an electrolyte and joined by a wire outside the electrolyte, a current was established in the circuit. The action causing such a difference of potential is called contact electromotive force and can be explained on the basis of the electron theory by that intrinsic property of metals known as electron affinity.

To understand this phenomenon, consider the configurations of the blocks of tungsten and nickel shown in Fig. 1. If the two blocks are placed as in part *a* and are in an uncharged state, there will be no difference of potential between them. Now, if they are brought together as in *b*, their contact surfaces must come to the same potential, but, when they are separated again to position *a*, they will exhibit a difference of potential and an electric field will exist between them. Joining the two blocks by a conductor as in *c* accomplishes the same result as moving them to position *b* and back to *a*. After the conductor connects the blocks in position *c*, a difference of potential exists between the blocks. The differences of potential mentioned here must be

* See frontispiece.

measured with apparatus that does not require current for its operation.

The reason for this strange phenomenon is as follows: The surface of the tungsten block has a potential energy barrier equal to its work function of 4.52 volts. Similarly, the nickel block has a potential energy barrier equivalent to 4.0 volts. When the two blocks are brought in contact as in *b*, the potential energy barriers must come to a common level. Since the nickel block has the lower electron affinity, electrons move from it across the boundary of contact into the bounding surface of tungsten and are held there by the image forces. When the blocks are separated the image charges are too far apart to have appreciable effect, the transferred electrons redistribute themselves over the blocks, and the nickel block is positive with respect to the tungsten. The difference of potential now existing is equal to the difference in the work functions, or $4.52 - 4.0 = 0.52$ volt, and is the *contact electromotive force* for these metals. In the action described the metal with the lower work function becomes positive, and hence this metal is said to be electropositive with respect to the other.

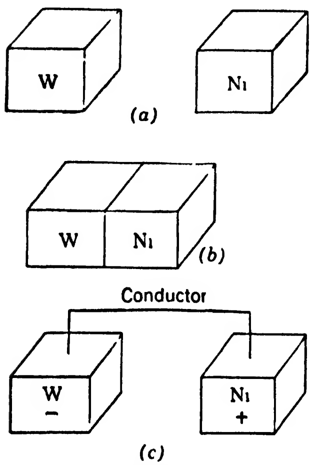


FIG. 1. Configurations illustrating contact electromotive force.

In the manufacture of electron tubes the electrodes are frequently made of different metals. These electrodes are usually joined by a circuit (conductor) similar to configuration *c* in Fig. 1. The contact emf's between these electrodes are added algebraically to any other differences of potential in these circuits. While these contact emf's are small in magnitude (usually a fraction of a volt), they may have a bearing on the operation of the tubes in which they exist when relatively low potentials exist on the electrodes. In high-mu triodes (such as 6Q7), which have a mu of approximately 100, the contact emf between the cathode and grid may be sufficient to provide the negative grid bias.*

Theory of Rectification in Vacuum. The theory of unilateral conductivity is simple and follows directly from the principles covered in

* Grid bias will be discussed in Chapter IV.

previous chapters. It will be assumed first that a hot electrode (cathode) and a cold electrode (anode) are placed in a tube where a good vacuum exists. The anode will be free (not connected to anything), as shown in Fig. 2. The battery A supplies current for heating the cathode, and when the optimum temperature is reached a cloud of electrons will surround the cathode. Each of the individual electrons in the cloud is thrown off with some initial velocity, and it moves out toward the anode against an attraction from its image positive charge on the filament. Since the anode is of neutral potential, it does not influence the electron. In all probability the kinetic energy of the electron due to initial velocity of emission will be overcome before it reaches the anode, and it will drop back to the filament. Now as point x is moved from a towards b, the current in the filament rises, which raises the temperature of the cathode, increases the rate of emission, and increases the initial velocity of emission of electrons. As this process is continued, a point will be reached where a few electrons will have a sufficient initial velocity to carry them over to the anode where they will "stick." The presence of these electrons on the anode will give it a negative charge, and this charge will now repel those electrons that approach the anode. Obviously, a point of equilibrium will soon be reached where no more electrons will become attached to the anode for the given temperature of the filament. If, now, point x is moved nearer to b, the initial velocity of some electrons will be raised so that they can overcome the small repulsion of the anode and land on it. Again, the negative potential on the anode will rise and a new condition of equilibrium will be reached. Since the anode is free and insulated electrically, the electrons that land on it have no avenue of escape.

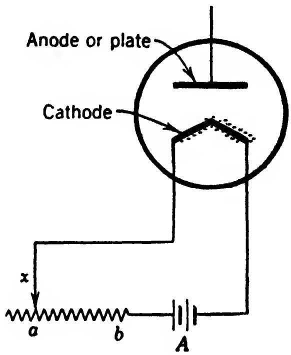


FIG. 2

If the circuit of Fig. 2 is changed by connecting the anode electrically to the filament as shown in Fig. 3, the action of the device is altered considerably.

The joining of the electrodes introduces a contact emf if different metals are involved. Assuming that the filament of the cathode is tungsten and the anode is nickel, a contact emf of 0.52 volt will exist with the anode positive with respect to the cathode. When the temperature of the cathode becomes sufficiently high, the initial velocity of

emission of some electrons will carry them to the anode as in the preceding case. This movement of electrons to the anode will be affected by the contact emf now present, and for the metals assumed the positive value on the anode will aid (attract) the passage of the electrons. Thus the electron transfer will be larger than in the preceding case. The electrons that land on the anode will not accumulate as before, since they are free to return to the cathode via the external conductor. Thus a very small continuous electron current will be produced by the

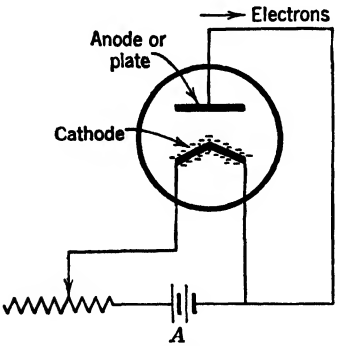


Fig. 3

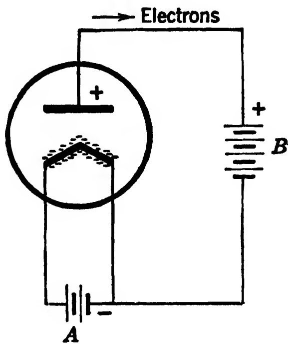


Fig. 4

electrons passing through the vacuum from cathode to anode and returning via the external path.

The addition of a battery in series with the anode circuit, as shown in Fig. 4, gives one common circuit for the use of the two-electrode tube. The theory of action becomes an extension of the preceding discussion. The hot cathode emits a cloud of electrons. Each electron that leaves the cathode is subject to a number of forces and factors that determine whether it moves over to the anode. The first factor is the initial velocity of emission determined by the temperature of the cathode. A second factor is the positive image charge on the cathode. A third and new factor is the attraction of the positive charge on the anode for the electron. A fourth factor is the negative space charge, which will be explained in the next article. The contact emf, if present, is a minor factor. The initial velocity of emission and the attraction of the anode are positive factors tending to carry the electron over to the anode. The attraction of the cathode is a negative factor and the negative space charge in general is also negative or opposing in its action. If the polarity of the *B* battery in Fig. 4 is reversed so that the anode becomes negative with respect to the filament, the anode will repel the

electrons that approach it and, except for a very low negative potential (fraction of a volt), no electrons will ever be able to land on the anode. Here the negative anode becomes the controlling factor and accounts for the unilateral conductivity of this type of vacuum tube. It should be noted that the positive potential placed on the hot cathode will not secure electrons from the negative anode because the anode is cold and is not a source of electron emission.

Negative Space Charge. The individual electrons emitted from a cathode are negative charges and as such they exert a force of attraction or repulsion upon surrounding charged bodies and particles. The cloud of electrons emitted from a filament thus acts as a charge or field in the region surrounding the cathode (Fig. 5). This cloud is very dense close to the filament and grows thinner rapidly as the distance from the filament is increased. The action of this space charge on an individually emitted electron is easy to analyze. As the electron breaks through the surface of the filament, it is repelled by the millions of other emitted electrons lying close to the filament as well as by those farther away. Thus at this point the space charge exerts a very powerful opposing effect upon this electron. As the electron moves out farther, as illustrated in Fig. 5, it is repelled toward the filament by all electrons in space between it and the anode and is repelled (aided) toward the anode by all electrons in space between it and the cathode. Obviously, the effect of space charge varies with the position of the electron under consideration. It is retarding for positions close to the cathode and aiding at positions nearer the anode.

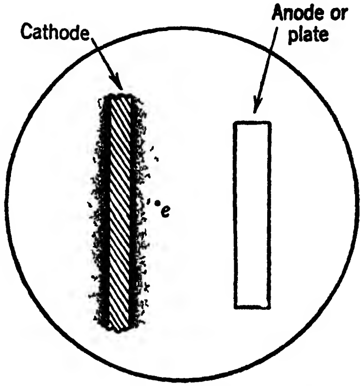


FIG. 5 Negative space charge surrounding a hot cathode.

The influence of negative space charge may be explained by means of potential distribution curves. Figure 6 (part *a*) illustrates a two-electrode tube having electrodes *C* and *A* with a difference of potential of 100 volts between them. Obviously, the potential at *C* is zero or ground potential, and, as the point of view moves from *C* to *A*, the potential must rise from zero to 100 volts. The manner of rise of potential from *C* to *A* may be shown graphically by curves as in part *b* of Fig. 6. If *C* and *A* consist of two large cold parallel plates, the

change of potential along a line near the center of the plates will be uniform (a constant potential gradient) and may be represented by the straight line *a*. If *C* is a small round wire (cold) and *A* is a hollow concentric cylinder surrounding *C*, the potential gradient near the wire will be high (owing to strong electrostatic field) and then will fall off as *A* is approached. Such a gradient may be represented by the dotted curve *b*. If *C* is a hot filament wire, the potential distribution may be a curve like *c*. Here the negative space charge surrounding the hot

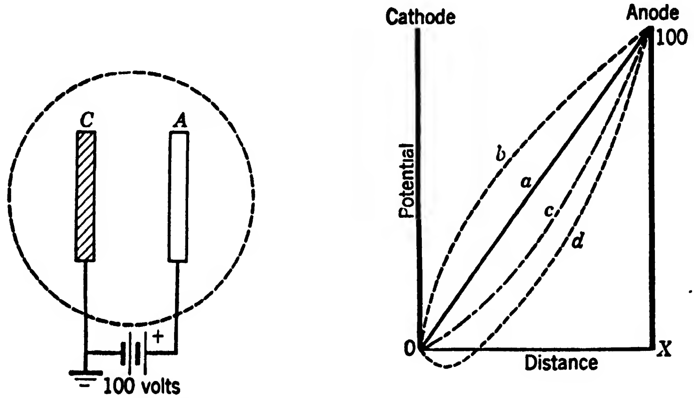


FIG. 6. Potential distribution in a two-electrode vacuum tube.

cathode depresses the potential in that region so that the curve dips below a straight line near *C*. If the cathode is heated to a high value, the electron emission and the space charge around *C* may become great enough to cause the potential in space near *C* to be actually lower than that of *C* and to give a potential distribution indicated by line *d*. If one imagines an electron being carried the distance from 0 to *X*, Fig. 6, curve *d*, it is easy to understand the depressing or opposing influence of the space charge near the cathode.

The negative space charge is very effective in limiting the number of electrons that pass to the anode of tubes having a high vacuum. It is so effective that it was once believed that there would not be any electron current in a perfect vacuum. Space charge can be overcome by high anode potentials and its effect can be neutralized by the presence of positive ions, as will be explained later.

Construction of Anodes. The anode is always the *electron-collecting electrode* in an electron tube. In the first tubes built, the anode was

constructed in the form of a little plate or sometimes two little plates connected in parallel and placed on each side of the filament type of cathode. The name plate has persisted down to the present time for the anodes of tubes used in communication circuits. For tubes used in power circuits and photoelectric devices, the term anode is more generally used. It is unfortunate that the term plate persists as applied to the anode because plate is a more fitting term for the electrostatic deflection electrodes used in cathode-ray tubes and similar devices. In describing tubes in this textbook, the author will attempt to use whichever term, *anode* or *plate,* has the most general acceptance in handbooks giving tube characteristics.

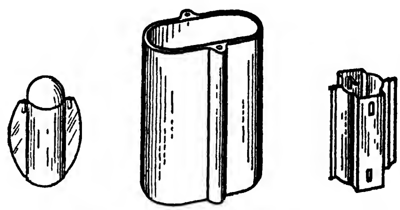


FIG. 7. Typical plate structures. (Courtesy Radio Corporation of America.)

The anode or plate is usually constructed so that it surrounds the cathode in the vacuum type of electron tube. This construction reduces the length of electron path and the resultant potential drop from cathode to plate. The typical plate structure is a hollow tube of circular, oval, or rectangular cross section, as illustrated in Fig. 7.

Many different materials have been used for constructing anodes. The desirable mechanical properties for this service are easy workability, mechanical strength at high temperatures, high thermal radiation, and low vapor pressure. Tungsten has desirable properties, but it is hard to form and hence has little use today for anode construction. Molybdenum has all desired properties except high heat emissivity. This weakness is overcome by use of radiating fins and a roughening of its surface and coating with materials such as zirconium powder (sintered).

Graphite has considerable use as an anode material. Its principal weakness is that it absorbs more gas than other materials. Nickel has a low melting point but can be used for low-power tubes. It is usually carbonized by a process wherein a well-adhering layer of amorphous carbon is deposited on the nickel. This process provides a heat emissivity approaching that of a black body and serves to reduce secondary emission. Tantalum is a satisfactory material for anode construction and is finding an increasing application in this field. Like molybdenum,

it usually requires fins on the anode and a roughened surface to increase the heat-radiating area.

The principal limitation in the design of an anode is its heat-radiating ability. The anode receives heat from two principal sources: (1) the heat radiated from the cathode and (2) the heat generated by the impact of the impinging electrons from the cathode.

If an electron is not intercepted in its flight from the cathode to anode, it strikes the anode with a kinetic energy equal to $\frac{1}{2}mv^2$. In high-vacuum tubes this kinetic energy is equal to the anode potential times the charge on the electron, or

$$e_b e = \frac{1}{2}mv^2$$

All this energy must be transformed into heat. Thus the power represented by anode voltage times the anode current is transformed into heat energy at the anode, and this energy must be dissipated by radiation, convection, and conduction to the outside medium. It is of interest to know that the bombardment of the anode by the electrons does eject numerous electrons from the anode through secondary emission. The electrons of secondary emission fall back into the anode and do not affect the magnitude of the anode current in the diode.

Characteristics of a Vacuum Diode. A two-electrode tube having a gas pressure of about 10^{-8} atmosphere or less is called a high-vacuum diode. A vacuum diode has two characteristics that are fundamental in understanding the operation of two-electrode tubes and the multi-electrode tubes to be covered in succeeding chapters. The two variables in the diode are the electron emission of the cathode and the potential applied between the cathode and the plate. If one is held constant and the other is changed, the two characteristics can be observed.

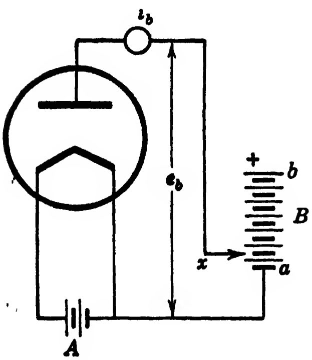


FIG. 8. Circuit for determining the voltage saturation characteristic of a vacuum diode.

The plate-voltage characteristic of a diode can be determined by connecting the tube in the circuit of Fig. 8. The filament or heater current should be held at rated value to give normal electron emission, and then the potential between the plate and cathode e_b can be varied from zero up to a high value. The resulting change of plate cur-

rent i_b with the variation in plate voltage is given in Fig. 9 for two different values of filament voltage. In both, the lower part of the curve rises rather slowly at first, then more rapidly for a time, and ultimately bends over and tends to flatten out or become horizontal. The slower rise of the plate current initially is caused by the strong retarding influence of negative space charge. For the lower curve the plate current begins to flatten out after 60 volts is reached. At this point the

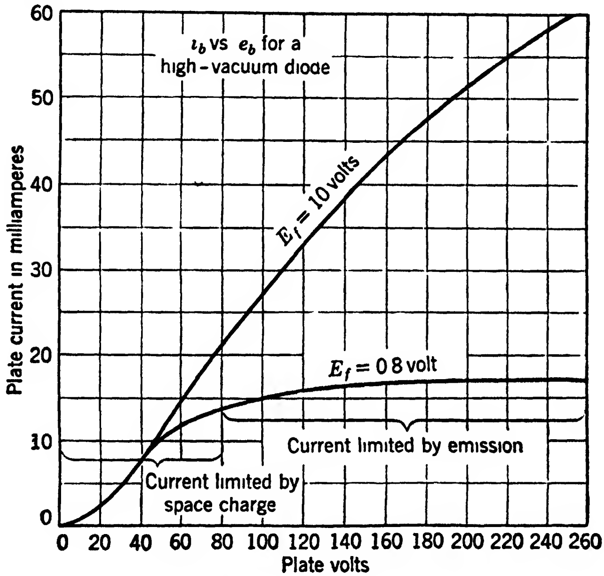


FIG. 9. Plate-voltage characteristic of a vacuum diode.

plate is attracting the majority of all electrons emitted, and a kind of "saturation" is rapidly approached as the plate voltage is raised further. If the filament temperature and resulting emission are raised by impressing a higher voltage on the filament (upper curve), saturation will be reached at much higher plate potential values. This is to be expected, for, with the emission of electrons greatly increased, the space-charge effect will be increased and also a higher plate voltage will be needed to attract the additional electrons before saturation is reached. The characteristic curves of a vacuum diode under discussion are frequently called voltage saturation curves.

The plate-current plate-voltage (i_be_b) characteristic curve for a vacuum diode is represented by an equation due to Childs.

$$i_b = A \frac{e_b^{3/2}}{x^2} \tag{1}$$

where i_b = total plate current.
e_b = total plate voltage.
x = distance between electrodes.
A = constant which depends on the geometry of the electrodes.

For a given tube where x is constant, the equation reduces to

$$i_b = ke_b^{3/2} \tag{2}$$

It should be noted that Childs' equation holds where all points of the cathode are at the same potential and at the same distance from the plate as in the heater type of cathode. This condition does not exist where the cathode is a filament having a varying potential along its length. Childs' equation also does not hold after saturation is approached or if electrons are emitted with an initial velocity.

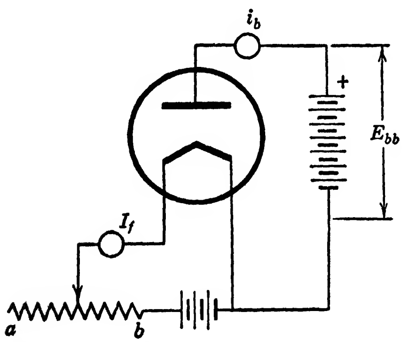


FIG. 10. Circuit for determining the temperature saturation characteristic of a vacuum diode.

The total plate current i_b covered by Childs' equation and the preceding discussion should never be confused with emission current. However, after saturation the plate current does represent the total emission current but not before.

A second characteristic of the diode is determined from the circuit of Fig. 10 by holding the plate supply potential E_{bb} constant and varying the filament current from zero up to the maximum permissible.† The variation of plate current for a series of plate supply potentials is illustrated in Fig. 11. The trend of these curves is approximately the same as those of Fig. 9 but the cause of the trends is reversed. The initial plate current is limited by emission, since the plate is attracting all the electrons being emitted. The trend of this part of the curve is similar to the curve that follows Richardson's equation. The i_b curve

* The exponent of e_b is $3/2$ for parallel-plane diodes. Other configurations require different exponents.

† See symbols for electron-tube circuits in front of book.

tends to flatten out at the higher filament currents because of the negative space-charge effect. Thus at low filament currents and temperatures the space charge is negligible and the plate attracts all the electrons emitted, but as the cathode emission rises the space-charge effect becomes stronger and stronger and finally prevents the plate

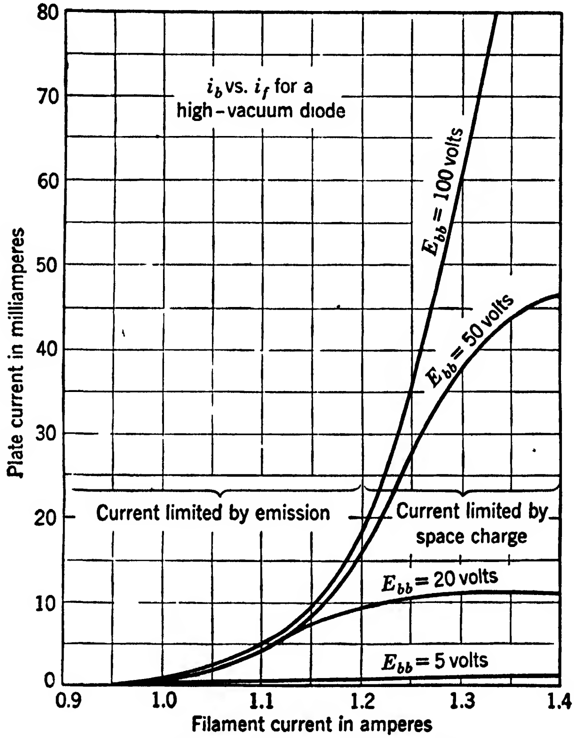


FIG. 11. Cathode-current characteristic of a vacuum diode.

from attracting additional electrons even though more are available in the expanding electron cloud. The curves of Fig. 11 are sometimes called temperature saturation curves. In these the plate attraction is constant with the negative space charge growing stronger, whereas in Fig. 9 the negative space-charge effect is constant and the plate attraction is growing stronger. In Fig. 11, the plate-current curve for $E_{bb} = 100$ volts rises more rapidly from zero because of the Schottky effect (page 35).

Rectification of Alternating Currents. The unilateral property of the two-electrode tube suggests a simple method of rectifying alternat-

ing current or transforming it to a unidirectional current. A two-electrode tube of the vacuum type is connected in a circuit as indicated in Fig. 12 with an alternating voltage impressed between the cathode and plate and a low-voltage a-c supply for the cathode heater. On the top half of the alternating-voltage cycle (part *b*) the plate is positive and electrons will pass to it. The electron flow will produce a current loop as indicated in part *c*. When the voltage drops to zero the current drops to zero, and when the plate becomes negative it repels the electrons emitted and no current results. Thus the alternating voltage results in a pulsating direct current which is termed half-wave rectification. If it is desired to utilize the idle portions of the alternating-voltage loops, a second rectifying tube may be added as shown in Fig. 13. An analysis of the operation of this circuit will show that, when the alternating-current supply main L_1 on the left is positive, the electrons move in the direction shown by the full-line arrows. Conversely, when the supply main L_2 on the right is positive, the electrons move as indicated by the dotted-line arrows. This action gives a full-wave rectification of current in the d-c load circuit as illustrated by part *c* of Fig. 13. This pulsating form of direct current is satisfactory for charging storage batteries and for electroplating. Whenever a steady d-c voltage or current is necessary, the pulsations can be smoothed out by filters as will be explained in Chapter XIV on rectifiers.

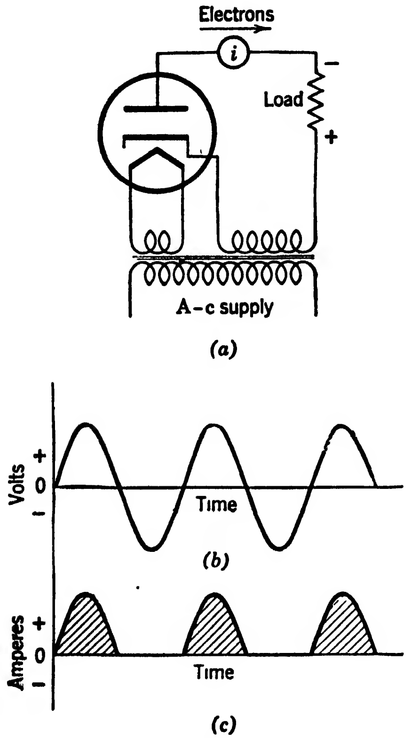


FIG. 12. Circuit and curves showing half-wave rectification of alternating current.

A simple circuit of a rectifier, a filter, and a potential divider is given in Fig. 14. This complete unit is called a power pack and is widely used in radio receivers, radio transmitters, and other electronic assemblies as the source of d-c voltage. The resistor at the right of

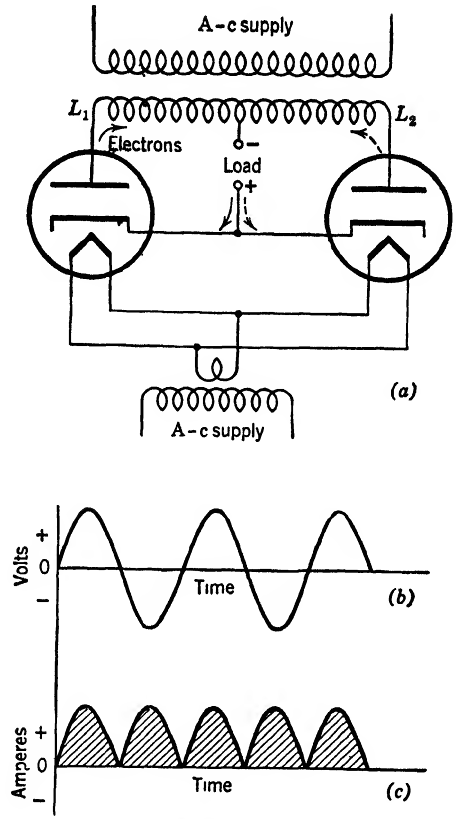


FIG. 13. Circuit and curves showing full-wave rectification of alternating current.

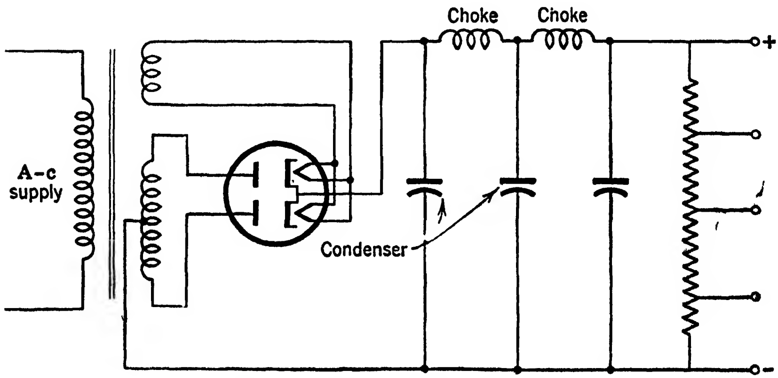
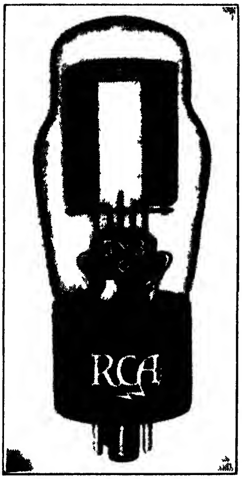
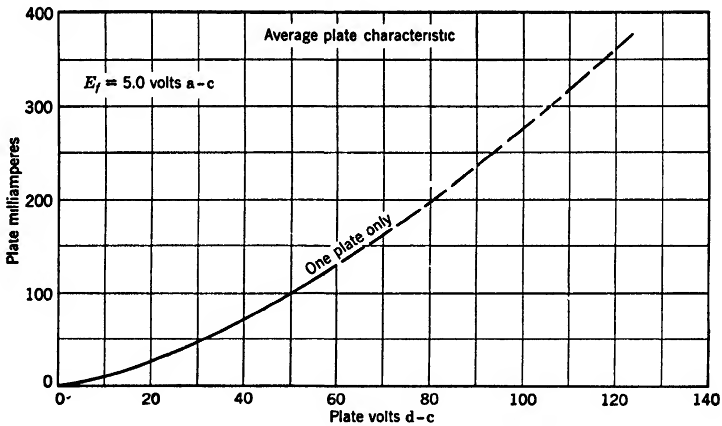


FIG. 14. Circuit for a power pack.



FULL-WAVE HIGH-VACUUM RECTIFIER

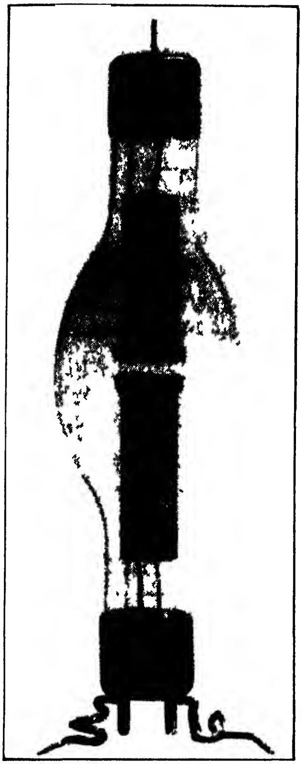
Filament—coated	
Voltage	5.0 a-c volts
Current	2.0 amp
Peak inverse voltage, max	1400 volts
Peak plate current per plate, max	375 ma
With Condenser-Input Filter	
A-c plate voltage per plate (rms), max	350 volts
Total effective plate-supply impedance per plate, min	50 ohms
D-c output current, max	125 ma
With Choke-Input Filter	
A-c plate voltage per plate (rms), max	500 volts
Input-choke inductance, min	5 henries
D-c output current, max	125 ma



TWIN DIODE

Heater—coated unipotential cathodes	
Voltage	6.3 volts
Current	0.3 amp
Direct interelectrode capacitances	
Plate 1 to cathode 1	3.0 $\mu\mu$f
Plate 2 to cathode 2	3.4 $\mu\mu$f
Plate 1 to plate 2, max	0.10 $\mu\mu$f
Rectifier	
A-c plate voltage, max	117.0 volts
D-c current, max	4.0 ma

FIG. 15. Vacuum rectifier tubes and characteristics. (Courtesy Radio Corporation of America.)



KENOTRON
WL-456

HIGH-VOLTAGE VACUUM RECTIFIER

General Characteristics

Air-cooled diode	
Filament voltage	11 volts
Filament current	20 amp
Filament heating time	30 sec
Net weight	2 lb
Shipping weight	6 lb

Maximum Ratings

Anode voltage, peak inverse	140,000 volts
Anode current, peak	500 ma
Anode current, average	60 ma

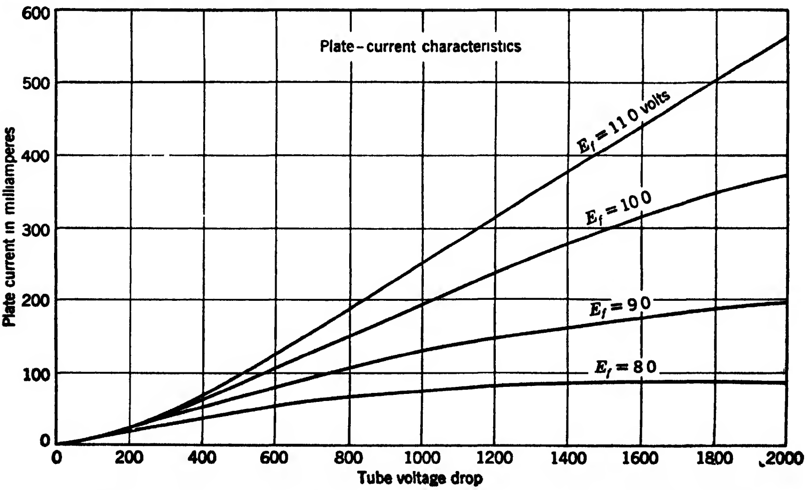


FIG. 16. Rating and characteristics of a high-voltage vacuum diode. (Courtesy Westinghouse Electric Corporation.)

the circuit with its numerous taps serves as a potential divider for furnishing a selection of voltages.

Applications of High-Vacuum Diodes. The applications of the vacuum diode are determined by its inherent properties. It is essentially a rectifier. Its plate current is limited by the negative space charge, so it has a small current rating. The potential drop from cathode to plate is rather high and varies with the plate current. This large voltage drop results in a low efficiency of rectification unless high voltages are used. In the low-voltage field the vacuum diode is used as a detector of radio signals and in power packs for radio receivers requiring currents of low magnitude. The characteristics of two tubes for this type of service are illustrated in Fig. 15. The vacuum diode will withstand a high inverse voltage, i.e., a high negative potential on the plate, without an arc-back. This makes it very desirable for high-voltage, low-current rectification. Thus it is widely used for furnishing direct current of the order of 1000 to 100,000 volts for X-ray machines, radio transmitting tubes, smoke precipitators, and dust eliminators. Vacuum diodes for the latter class of service are known as *kenotrons*. A typical kenotron and its rating and characteristics are illustrated in Fig. 16.

Problems

1. A certain vacuum diode has an emission of 10 milliamperes for a plate potential of 50 volts. Assuming that Childs' law holds for the tube, calculate and plot the saturation-emission current for 20-volt steps from zero to 100 volts on the plate.

2. Assume in Fig. 16 a cathode-anode separation of 2 centimeters and a peak applied potential of 20,000 volts. What will be the velocity of an electron when it hits the anode if it left the cathode with zero initial velocity? What will be the energy of this electron on impact in (a) electron volts? (b) joules?

3. What is the average internal or plate resistance of the upper vacuum diode shown in Fig. 15?

4. Calculate Problem 3 for the kenotron of Fig. 16, when $E_f = 11$ volts.

5. The kenotron of Fig. 16 is employed for half-wave rectification, feeding a load resistor of 10,000 ohms. What is the voltage across the load resistor when the drop across the tube is 1000 volts? What power is being dissipated at this instant in the load? within the tube (including filament power)?

REFERENCE

McArthur, E. D., "Electronics and Electric Tubes," *Gen. Elec. Rev.*, December 1933.

Chapter IV

GRID-CONTROLLED VACUUM TUBES

Three-Electrode Vacuum Tube. In 1907 De Forest added a third electrode to the Fleming valve and called the new device the audion. The third electrode, which he called a grid, consisted of a zigzag wire which was placed in a vacuum tube between a heated filament and a plate. The value of the grid lay in its ability to control the electron current between the cathode and the plate, and this invention constituted the most important development of the twentieth century. It has extended the field of communication by wire across continents, it has made possible radio communication around the world, and it is now revolutionizing the use and control of electric power in the industrial world.

The modern three-electrode vacuum tube, known by the family name of triode, uses cathodes and plates like those described in the preceding chapters. The grid is usually a coil of fine wire wound in the form of a helix and interposed between the cathode and plate. The construction may consist of circular concentric cylinders as shown schematically in part *a* of Fig. 1, or it may utilize oval or oblong cylinders as in part *c* of the same figure. The standard symbols for the triode are given in part *b* of Fig. 1; the upper symbol represents the filament type of cathode and the lower the heater type. The cathode heater is not an electrode.

Theory of Grid Action. The grid of the three-electrode vacuum tube functions by a change of the charge residing upon the grid. This change in charge (and potential) serves to control the electron stream between the cathode and the plate. The process of control can be visualized in a number of different ways. One simple visualization follows from Fig. 2, which is a schematic diagram of a cross section taken through the axis of a three-electrode tube. Here the helical wire grid appears as circles. When the cathode is heated it will emit a cloud of electrons, part of which will be attracted over to the plate by its positive charge. The electrons that go to the plate must pass through the meshes of the grid and hence will be affected by the potential residing

on the grid. Assuming that the grid is positively charged with respect to the cathode, an electron in position *a* will be subject to several forces acting on it: first, the initial velocity of emission; second, the

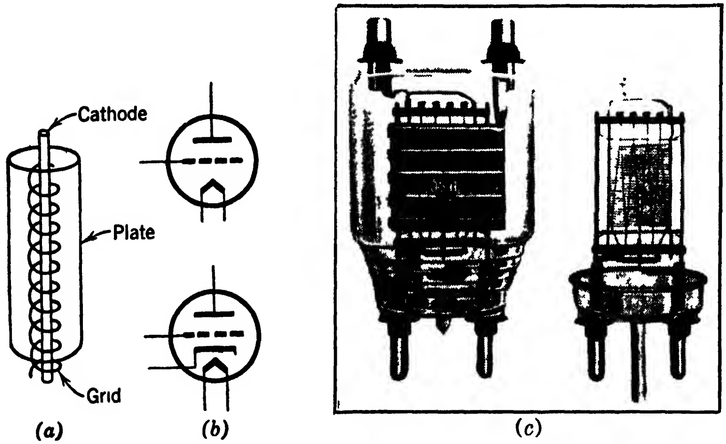


FIG. 1. (*a*) Schematic triode. (*b*) Triode symbols. (*c*) Triode parts and assembly. (Courtesy Radio Corporation of America.)

attraction back to the cathode due to the image positive charge left on the cathode; third, the influence of negative space charge; fourth, the attraction due to the plate; and, last, the attraction of the charge on the grid wires. The attraction of the grid wires will be in the directions indicated, but the resultant of all these forces will be toward the plate. Thus the grid aids (controls) the passage of electrons to the plate. If an electron progresses to the point shown at *b*, it will be subject primarily to three influences: (1) its instantaneous velocity, (2) the attraction of the plate, and (3) the resultant attraction of the grid wires. This resultant force of the grid wires is zero, and hence the grid is now ineffective, but it has accomplished its task in helping the electron to escape from the cathode and the repulsion of negative space charge. An electron in the position *c* is subject to the same group of forces as any other electron, but the resultant of the attraction of the grid wires and the pull of the plate potential is directly toward the grid wire, and

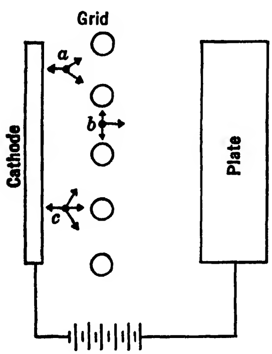


FIG. 2. Forces acting upon electrons in a triode.

hence this electron will land on the grid. Electrons moving on a line passing to one side of the grid wires will probably pass to the plate because of their high velocity and because the grid potentials are usually low relative to the plate potential.

If the potential of the grid is now made negative with respect to the cathode and a like process of reasoning is applied, it is evident that the action of the grid will always oppose the passage of the electrons to the plate. The degree of opposition offered will depend on the magnitude of the negative potential. It is apparent that a strong negative po-

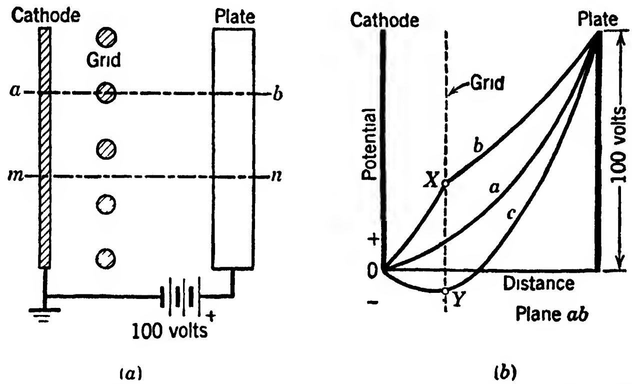


Fig. 3 Potential distribution in a triode (plane through grid).

tential or negative bias may bar all electrons from passing to the plate or from passing to the grid itself. If the grid is given a small negative bias, changes of potential superimposed on this bias will produce corresponding swings in the cathode-plate current.

A second method of analyzing grid action involves a study of the potential distribution in the three-electrode tube. If in Fig. 3, part *a*, a plane *ab* is passed perpendicularly to the axis of the tube, it will show a potential distribution from cathode to plate as in part *b*. Let it be assumed that the cathode is at zero or ground potential and that the plate is maintained at a potential of 100 volts above the cathode by the battery. If the grid were omitted and the cathode heated, the potential distribution would be given by curve *a* of Fig. 3*b*, the depression of the curve being due to the negative space charge. Now, if the grid is added and raised to a small positive potential, the potential distribution will be raised to that of curve *b* with point *X* at the exact potential of the grid. Again, if a negative potential is applied to the grid, the potential distribution will fall to that of curve *c*, where *Y* is the negative potential applied to the grid. The change in the potential

distribution will not be so pronounced for other planes in the three-electrode tube. Thus if a plane is passed through *mn* of Fig. 3, part *a*, the potential distribution will be represented by curves *a′*, *b′*, and *c′* of Fig. 4. Planes passed through intermediate points would show potential distribution curves varying between the limiting cases illustrated in Figs. 3 and 4. (In any case the effect of the potential on the grid is obvious. When the grid potential is raised, it aids the electrons in breaking away from the cathode and moving to the plate, whereas a negative potential opposes the electron movement to the plate.)

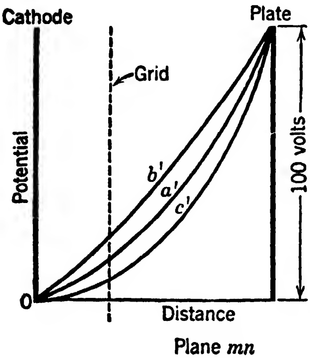


Fig. 4. Potential distribution in a triode (plane between grid wires).

It is evident from the preceding paragraph that the closeness of the wires in the grid will affect the degree of the grid control of the electron stream. If a coarse mesh is used, electrons will pass midway between the wires with a relatively small influence from the grid. A fine or close mesh will give a more uniform and effective control. Some tubes have a coarse grid near the center and a fine mesh at the top and bottom. These are known as variable-mu tubes and their advantage will be pointed out later.

A third method of visualizing grid action in a triode is through a study of patterns in the electric field in the cathode-plate space under a variation of the potential applied to the grid. Pictures of possible electric fields are given in Fig. 5. Part *a* of this figure shows a field of uniform direction and distribution which would exist without any grid (diode). In part *b*, it is assumed that a grid has been inserted and its potential has been made the same as the potential of the space it occupies. Under these conditions the grid does not influence the field. In part *c*, the grid has been connected to the cathode so that it is at zero or cathode potential. Since the grid is closer to the plate than the cathode is, many of the lines of the electrostatic field now emanate from grid to plate. In part *d*, the grid has been made slightly positive with respect to the cathode. Here the field lines from the negative cathode lead to both grid and plate. This appears the same as in part *b* where the grid was positive, although there may be a difference in the magnitude of the field in the cathode-grid space. A change of the grid potential to negative in part *e* of Fig. 5 causes a reversal in the direction of some lines in the electric field. Now, some of the field

lines from grid lead to the cathode. The pictures of the electric field are completed by part *f* where the grid is made strongly negative so that all field lines to the cathode have been reversed. An analysis of the action of these various patterns of electric-field direction and distribution upon electrons released at the cathode will show the effectiveness of grid control.

The early triodes and most of those in use today have the grid located between the cathode and the plate. It is possible to place the grid

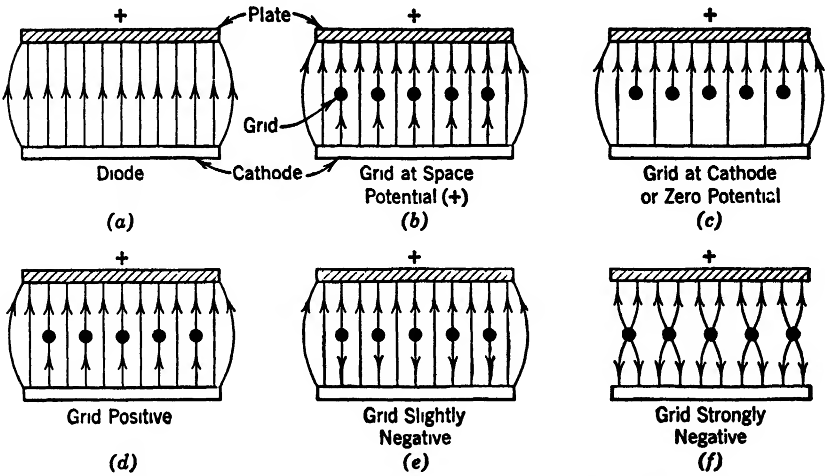


FIG. 5. Electric field distribution in triodes under different grid potentials.

in other positions. Thus one type of tube has a central rod for a grid, which is surrounded by a helical coil (cathode) with the plate outside the cathode. This is known as an internal-grid tube. The control of the grid is not so effective, but the closeness of the cathode to the plate results in a relatively large cathode-plate current, with the expenditure of little or no power in the grid circuit. This tube is useful for certain types of relay work, oscillograph amplifiers, and speech-frequency amplifiers.

A conventional circuit for the three-electrode tube using batteries for the necessary applied potentials is shown in Fig. 6. The filament type of cathode is heated by a battery called the *A* battery, the plate secures its source of potential from another battery known as a *B* battery, while the input or potential variation to the grid is connected between the grid and the cathode. If the grid is to be kept normally negative, a battery called a *C* or biasing battery is connected into the grid supply circuit. The voltages supplied by these three batteries are

represented by the symbols E_{ff}, E_{bb}, and E_{cc} as shown in Fig. 6. If at any time the grid becomes positive with respect to the cathode, it will attract a few electrons to itself, and these electrons will return to the cathode via the grid-cathode circuit. These electrons constitute the grid current. The plate current is usually several times as large as the grid current when the grid is slightly positive. When the grid is nega-

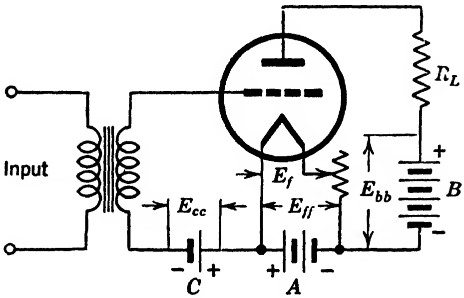


Fig. 6. Conventional circuit for a triode.

tive, it is generally assumed that the grid circuit does not carry any current, although this is not rigidly true as will be shown later.

A *free grid* is one that is isolated from any circuit through which a continuous stream of electrons may pass. This may mean that the grid is entirely free as in part *a* of Fig. 7, or it may be connected to a series circuit containing a condenser which will block any continuous

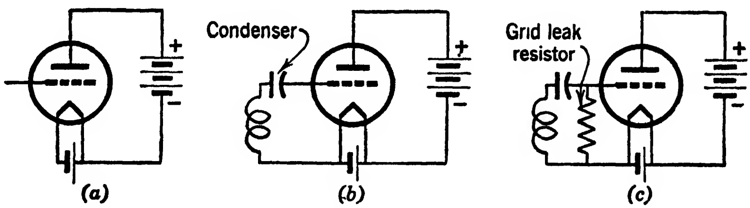


Fig. 7. Free-grid and grid-leak circuits.

flow of electrons in either direction (part *b*, Fig. 7). Before a triode containing a free grid is energized, the grid will probably be at zero or ground potential. As soon as the cathode and plate are energized (part *a* or *b* of Fig. 7), electrons from the cathode may move out under their initial velocity of emission and attraction of the plate and land on the grid. Since these electrons cannot leave the grid through any external conductor, the grid will assume a negative potential of such a magnitude as to bar any more electrons from landing on it. Also, if a

varying signal voltage is impressed on the grid through a condenser, it will provide transient changes in grid bias which may permit the grid to go even more negative until it serves as a complete block to any electron movement to the plate. The accumulation of a strong negative charge on a free grid makes it impossible to use the triode in any useful circuit. Accordingly, it is necessary to provide some path for the accumulating charge to leak back to the cathode. In some circuits the input circuit itself will provide a natural leak for the grid. In other circuits where the input to the grid contains a condenser as in part *b* of Fig. 7, it is necessary to provide a high-resistance path called a *grid leak* as illustrated in part *c* of Fig. 7. A grid leak has a value from one-half up to several megohms.

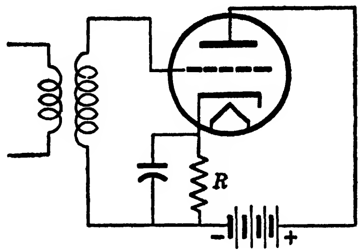


FIG. 8. Circuit for producing negative grid bias.

The triode and the multielectrode tubes to be described later are operated with the grid held negative with respect to the cathode. This negative grid or bias prevents any normal electron current from cathode to grid which would represent a useless current and power loss, and it also provides better operating characteristics in the tubes themselves. The grid bias may be provided by three methods. One method using a *C* battery is illustrated in Fig. 6. A second method provides the desired bias by the selection of a suitable value for the grid leak shown in part *c* of Fig. 7. The leak of electrons back to the cathode through this high resistance provides an *RI* drop with the grid held negative. A third method for obtaining a negative bias utilizes a voltage drop arising from the direct current in the cathode-plate circuit. This is illustrated in Fig. 8 where the current through *R* provides the desired potential drop. The condenser in parallel with *R* provides a low-impedance shunt path for any signal frequencies in both the grid and plate-cathode circuit. In a few applications a low grid bias is provided by the contact emf between the cathode and the grid.

Characteristics of a Three-Electrode Tube. The important characteristics of the three-electrode tube shows how the anode or plate current changes when either the grid voltage or the plate voltage is varied, with the other held constant. The transfer characteristic of a triode may be determined through the use of the circuit given in Fig. 9. The cathode is heated to its normal operating temperature and a

voltage E_{bb} is applied to the plate. The grid is supplied by a circuit for varying the potential impressed upon it from a range of negative values up through a series of positive values. This variation is secured by moving the point X along the potentiometer from point a to point b. At point a the grid may be made so strongly negative that no elec-

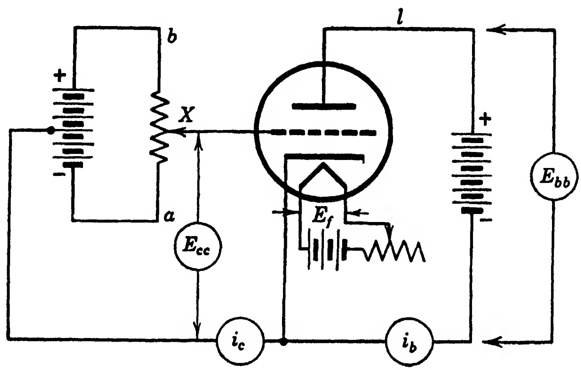


FIG. 9. Circuit for determining a triode transfer characteristic.

trons can pass to the plate. As the grid is made less negative, some electrons do pass to the plate, and with the movement of X to b the plate current i_b rises along the curve as shown in Fig. 10. After the grid becomes positive with respect to the cathode, a small current i_c begins to pass to the grid and follows the trend shown by the dotted curve on Fig. 10. It should be understood that the magnitude of the grid current is generally smaller than that indicated. The curve of Fig. 10 shows the operation of a tube for only one value of plate voltage, and, since a range of positive potentials may be used on the plate, the complete picture of the characteristic of a tube must be obtained through a family of curves for different plate potentials, as shown in Fig. 11 (lower right). A second set of family curves given in the upper right view of the same figure shows how the plate current varies with changes in the plate voltage. These curves are known as the plate characteristic curves, and they are very useful in designing electronic circuits. The plate characteristic curves may

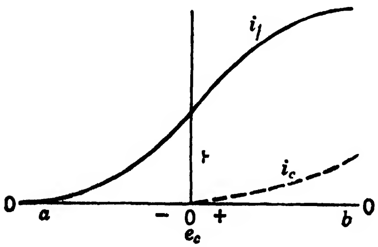
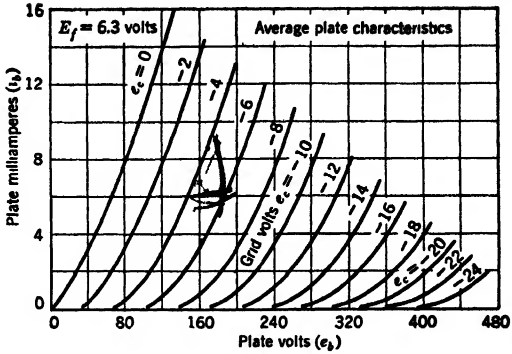


FIG. 10. Transfer characteristic of a triode.



DETECTOR-AMPLIFIER TRIODE

Heater—coated unipotential cathode	
Voltage	6 3 volts
Current	0 3 amp
Direct interelectrode capacitance	
Grid to plate	3 4 $\mu\mu$f
Grid to cathode	3.4 $\mu\mu$f
Plate to cathode	3 6 $\mu\mu$f
Maximum overall length	2⅝ in
Maximum seated height	2$\frac{1}{16}$ in
Maximum diameter	1$\frac{5}{16}$ in

Amplifier

Plate voltage, max	300 volts
Grid voltage, min	0 volts
Plate dissipation, max	2 5 watts
D-c heater-cathode potential, max	90 volts
Cathode current, max	20 ma

Typical Operation—Class A_1 Amplifier

Plate	90	250 volts
Grid	0	−8 volts
Amplification factor	20	20
Plate resistance	6700	7700 ohms
Transconductance	3000	2600 μmhos
Plate current	10	9 ma

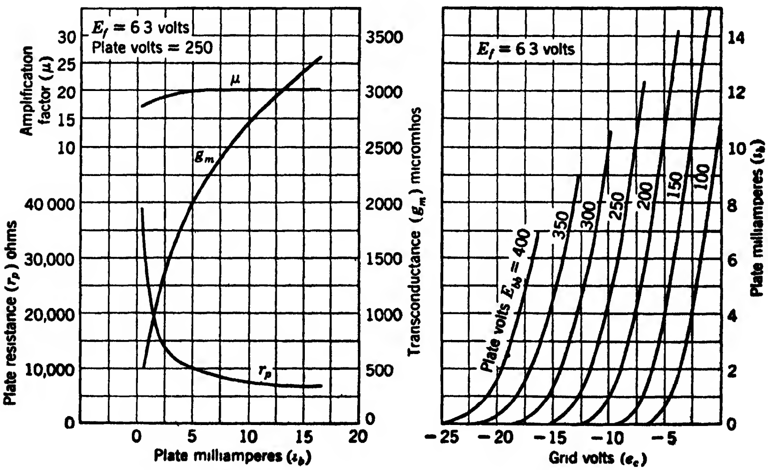


FIG. 11 Rating and characteristics of a detector-amplifier triode. (Courtesy Radio Corporation of America.)

be obtained directly by using a suitable test circuit, or they can be obtained by replotting the data given in the transfer characteristics.

Both sets of curves covered in the preceding paragraph are static characteristics, i.e, they apply for a static or constant potential between the cathode and the plate. In the application of the three-electrode

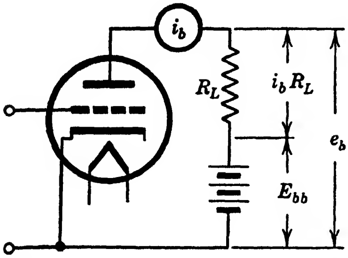


FIG. 12. Load circuit on a triode.

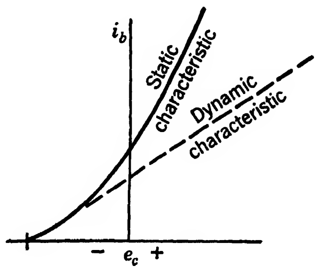


FIG. 13. Static and dynamic characteristic of a triode.

tube some form of load must be placed in the plate circuit. This load (Fig. 12) will present a resistance R_L to the current and will give a fall of potential over itself. Thus the potential between the cathode and plate does not remain constant as E_{bb} but becomes e_b where $e_b = E_{bb} - i_bR_L$. This fall in plate potential will reduce the plate current

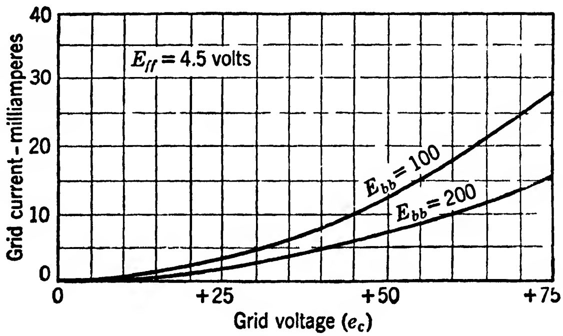


FIG. 14. Typical grid-current characteristics of a triode.

below the values determined by the circuit of Fig. 9. The effect of this lowering of plate potential with the increase of grid potential is to give a lower curve as shown dotted in Fig. 13. This new curve is the dynamic characteristic for the given load resistance. Obviously, the dynamic characteristics for a known load resistance can be determined

experimentally by inserting this load at point l on Fig. 9. However, since there are an infinite number of possible load resistances, it is not feasible to give all dynamic characteristic curves. Hence it is customary to determine the dynamic curve for a given load from a family of static curves by suitable calculations.

Two grid-current characteristics are given in Fig. 14. With a higher plate voltage E_{bb}, the plate attraction is relatively stronger and fewer electrons land on the grid, resulting in smaller values of grid current.

Amplification Factor. The chief value of the vacuum type of triode lies in its ability to amplify a relatively weak signal in the form of a change of potential or charge impressed across its grid. Such a change of potential will produce a rather large change of current in the plate circuit, and this change of current passing through a resistance or one winding of a transformer will produce a change of voltage of increased magnitude. It is possible for a mere change of charge on the grid (representing zero or nearly zero power input) to produce a large change of current and voltage in the plate circuit. Thus there would appear to be an infinite increase of power. In actual circuits some power is absorbed in the input circuits so that infinite amplification of power is not attained.

The amplifying power of a triode is the measure of the greater effectiveness of changes in the grid potential over those of the plate potential. The obvious way of stating such a measure is by means of a ratio of the voltages employed. Thus, mathematically, the amplification factor mu (μ) may be expressed as

$$\mu = -\frac{e_b - e_b'}{e_c - e_c'} \quad (\text{for } i_b = \text{constant}) \tag{1}$$

where $e_b - e_b'$ is the change in plate voltage required to compensate for a small change in grid voltage represented by $e_c - e_c'$. Since a decrease in plate voltage is necessary to compensate for an increase in grid voltage, the minus sign is necessary if mu is to be considered as a positive number. The factor mu may be determined from the transfer characteristics of the tube by substituting values in equation 1. Referring to Fig. 11 (lower right), it will be found that 11 milliamperes of plate current will be produced by 100 volts on the plate and a 0.0 volt on the grid. This same current will be produced by 200 volts on the plate and -5 volts on the grid. Thus

$$\mu = -\frac{(100 - 200)}{0 - (-5)} = \frac{100}{5} = 20$$

Similar calculations for other values of constant current will give a like value for mu. If points are taken for low values of plate current where the curvature of the characteristic is high, the value calculated for mu will vary.

In terms of calculus, the amplification factor is expressed as follows:

$$\mu = -\frac{\partial e_b}{\partial e_c} = -\frac{de_b}{de_c} \quad (i_b = \text{constant}) \tag{2}$$

The amplification factor depends on the geometry of the tube. The closer the control grid can be placed to the cathode, the more effective

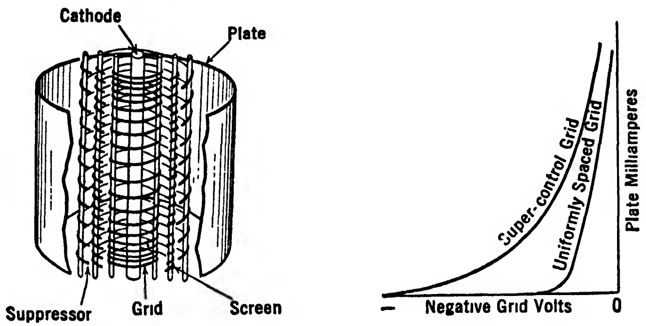


FIG. 15. Construction and curves of a variable-mu tube. (Courtesy Radio Corporation of America.)

it will be in controlling the electron stream to the plate, and hence the higher the amplification factor will be. This closeness is limited by the necessary insulation of parts as well as by mechanical strength and construction. The amplification factor of a triode varies from about 3 as a minimum to about 100 as a practical maximum, the exact value depending upon the purpose for which the tube was designed.

For some applications it is desired to have a variable mu. This is secured by a close spacing of the grid wires at the top and bottom of the grid and a coarse spacing at the center. This construction and the effect upon the plate-current characteristic are illustrated in Fig. 15. The characteristic curve of a triode and other multielectrode tubes near the cutoff point is broadened, and the tube can be used for amplification at lower levels of plate current. This type of construction is known by the terms supercontrol grid, remote cutoff, and variable mu. Tubes employing this type of grid are used for automatic volume control.

Mutual Conductance and Transconductance. The mutual conductance of a triode is the rate at which the plate current changes with the grid voltage. Mathematically, it is the derivative of the plate current with respect to the grid voltage.

$$g_m = \frac{\partial i_b}{\partial e_c} = \frac{di_b}{de_c} \quad (e_b = \text{constant}) \tag{3}$$

Mutual conductance (g_m) is the slope of the transfer characteristic curve and is expressed in terms of microamperes per volt, or micromhos. The concept of mutual conductance may not be easy to grasp. It may help to think of the conductance of (one circuit) the plate-cathode circuit as being affected (mutually) by the voltage impressed across (a second circuit) the grid-to-cathode circuit.

For an example, take the point on the lower right-hand curves of Fig. 11 corresponding to 5 milliamperes on the 100-volt plate curve. Assume a current swing from 4.0 to 6.0 milliamperes which corresponds to a change of 1.0 volt on the grid. Hence

$$g_m = \frac{2 \times 10^3}{1} = 2000 \text{ micromhos}$$

Since the adoption of multigrids or electrodes for tubes, the term transconductance often displaces the expression mutual conductance. The mutual conductance referred to above becomes the control-grid-to-plate transconductance, and similar terms designate the electrodes involved in the transconductance in multielectrode tubes. The value of the transconductance varies because it depends on the varying curvature of the transfer characteristic of a tube.

Plate Resistance. The plate resistance of a triode is the rate of change of plate voltage with respect to the rate of change of plate current. Mathematically, it is the derivative of the plate voltage with respect to the plate current.

$$r_p = \frac{\partial e_b}{\partial i_b} = \frac{de_b}{di_b} \quad (e_c = \text{constant}) \tag{4}$$

Plate resistance is the reciprocal of the slope of the plate characteristic curves (Fig. 11, upper right). This plate resistance is frequently referred to as the a-c resistance or the dynamic resistance to distinguish it from the d-c plate resistance. The d-c plate resistance is the quotient of the static plate to cathode voltage E_{bo} divided by the static or quiescent plate current I_{bo}. The d-c plate resistance is of little im-

portance in tube operation. For a sample calculation of a-c plate resistance r_p, use the curve $e_c = -4.0$ and $i_b = 10$ milliamperes for the upper right curves of Fig. 11. Here a swing of 40 volts on the plate will give a current change of 5.3 milliamperes. Thus

$$r_p = \frac{40}{5.3 \times 10^{-3}} = 7500 \text{ ohms} \quad \text{(approx)}$$

Parameters of Multielectrode Vacuum Tubes. The three factors of amplification—constant μ, control-grid-to-plate transconductance g_m, and plate resistance r_p—are termed the parameters of multielectrode vacuum tubes. These factors are related to each other as follows:

$$r_p \times g_m = \mu \tag{5}$$

This relationship is proved by a substitution of the expression for each of these three factors from equations 2, 3, and 4.

$$\frac{\partial e_b}{\partial i_b} \times \frac{\partial i_b}{\partial e_c} = \frac{\partial e_b}{\partial e_c}$$

The magnitude of the parameters depends on the geometry of the tube, such as the spacing of the cathode, grid, and plate, the diameter of the grid wires, the spacing of the grid wires, and the area of the plate. The parameters of a tube are important in the design of vacuum-tube circuits.

The parameters of a triode vary with plate current as illustrated in Fig. 11 (lower left) for a 6J5 triode. The amplification factor is usually constant throughout a wide range of plate current, but the transconductance and plate resistance usually show much variation. In the design of circuits the values of the parameters are obtained from data furnished in the tube manuals of the manufacturers, or they may be calculated from test data using the methods suggested in the preceding discussion.

Plate Current in a Triode. Childs' equation for the two-electrode tube can be modified slightly for application to the triode. The modified equation becomes

$$i_b = K\left(\frac{e_b}{\mu} + e_c\right)^{3/2} \tag{6}$$

where K is a constant depending on the tube dimensions. Obviously, the expression in the parenthesis refers to the grid potential and not to the plate potential of Childs' equation. This equation does not hold for low values where mu is not constant nor for high values where

saturation effects are present; also, for positive values of grid potential, i_b represents the sum of plate and grid currents.

Types of Triodes. Triodes may be classified on the basis of (1) their construction, (2) their use, or (3) their power rating. Under the first classification triodes are built with all glass enclosures, in metal envelopes with air cooling, and in metal tubes with water cooling. Most tubes use glass enclosures. Glass construction gives a lower cost because it utilizes the manufacturing technique developed through years in producing electric light bulbs. Glass has a low heat-dissipating ability and is limited to use in tubes having a maximum plate dissipation of approximately 1000 watts. Two kinds of glass are used in forming glass tubes—an ordinary soft glass for small tubes and for tubes of low rated capacity, and a hard glass having a higher softening and melting temperature for power tubes, which must dissipate considerable heat and operate at a high temperature.

Two glass triodes which have wide use as amplifiers for repeaters on telephone toll lines are shown in Fig. 16. A low-power triode having a glass enclosure and a graphite plate is illustrated in Fig. 17. Its moderately high grid- and plate-voltage rating makes it adaptable for use in transmitter circuits as an amplifier, oscillator, or modulator. Interesting glass triodes are shown in Fig. 18. These tubes are very small and compact and are designed for ultra-high-frequency service. For this service it is necessary to reduce the interelectrode capacities to a minimum, which means small electrodes and short lead-in conductors.

The metal receiving tube is small in physical size and its metal case serves as a shield from external fields. The active parts of the tube are of normal size, but a large saving in space is made in the stem (see Fig. 19) and in the metal envelope which fits the electrode assembly closely. The metal tube utilizes a new alloy called Fernico and a glass which have practically the same coefficient of expansion throughout a wide range of temperature. The Fernico base of the metal tube consists of a flanged disk in which a number of eyelet holes are pressed. Glass beads are welded into these eyelets so as to support the leads and insulate them from each other and the metal case (Fig. 19). A small metal tube is welded into the center of the disk and serves as the medium for exhausting the assembled electron tube. After the electrodes are assembled on the base, the metal cover is welded to the base in a single quick butt weld. Then the tube is exhausted and a pinch weld on the exhaust tube seals the triode.

The amount of heat energy that can be dissipated by radiation and convection in air from glass and metal tubes is rather limited because

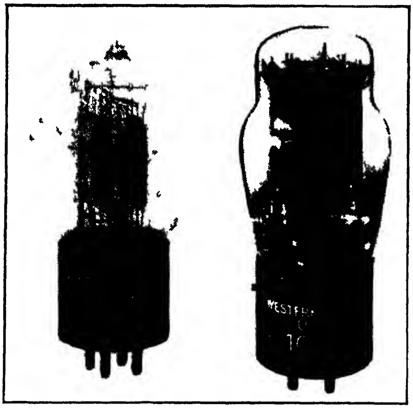


FIG 16 Two filamentary triodes for telephone service *left* voltage-amplifier, *right,* low-power amplifier for voice and carrier frequency (Courtesy Western Electric Company)

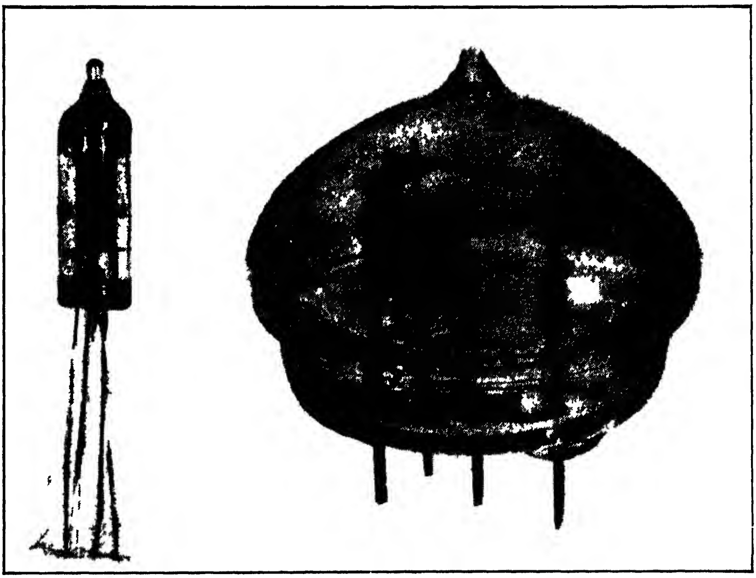
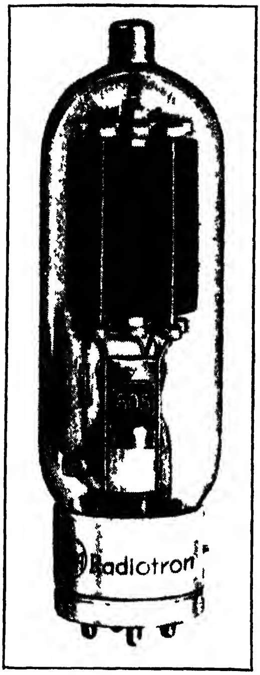


FIG 18 Ultra-high-frequency amplifier and oscillator triodes *left,* 6K4 (courtesy Sylvania Electric Products, Inc), *right,* 316A (courtesy Western Electric Company).



R-F POWER AMPLIFIER AND OSCILLATOR

General Data

Filament—thoriated tungsten	
Voltage (a-c or d-c)	10 volts
Current	3.25 amp
Direct interelectrode capacitances (approx)	
Grid to plate	6.5 $\mu\mu$f
Grid to filament	8.5 $\mu\mu$f
Plate to filament	10.5 $\mu\mu$f

Class C Telegraphy

D-c plate voltage, max	1500 volts
D-c grid voltage, max	−500 volts
D-c plate current, max	210 ma
D-c grid current, max	70 ma
Plate input, max	315 watts
Plate dissipation, max	125 watts

Typical Operation

Filament voltage	10	10	10 a-c volts
D-c plate voltage	1000	1250	1500 volts
D-c grid voltage	−95	−100	−105 volts
Peak r-f grid voltage	225	230	235 volts
D-c plate current	200	200	200 ma
D-c grid current, approx	40	40	40 ma
Driving power, approx	8.5	8.5	8.5 watts
Power output, approx	130	170	215 watts

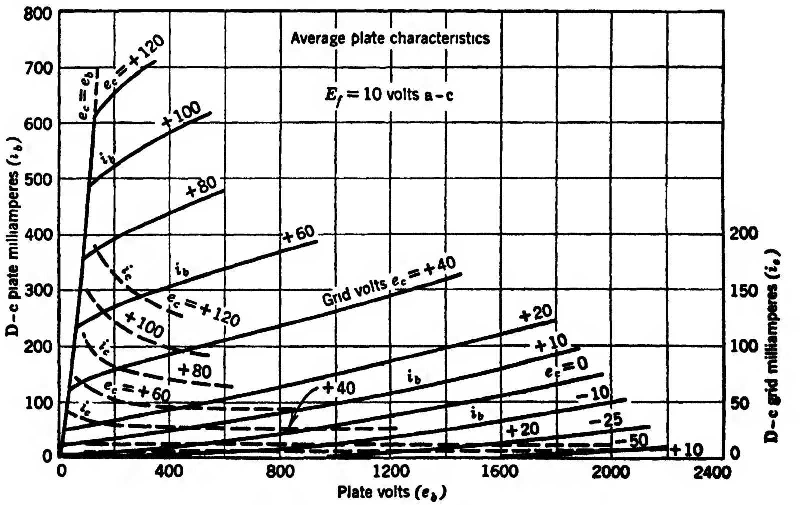


FIG. 17. Rating and characteristics of a low-power vacuum triode. (Courtesy Radio Corporation of America.)

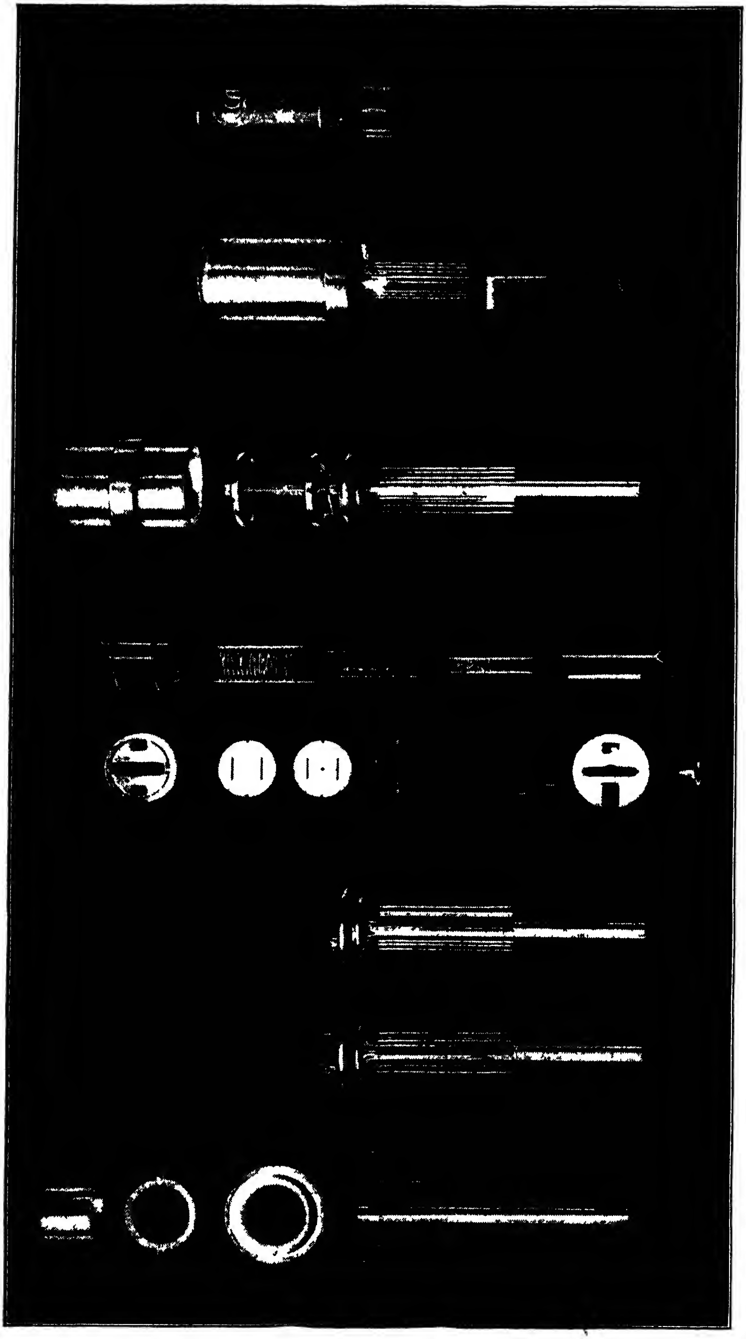


FIG 19. Parts and assembly of a metal receiving tube. (Courtesy Radio Corporation of America)

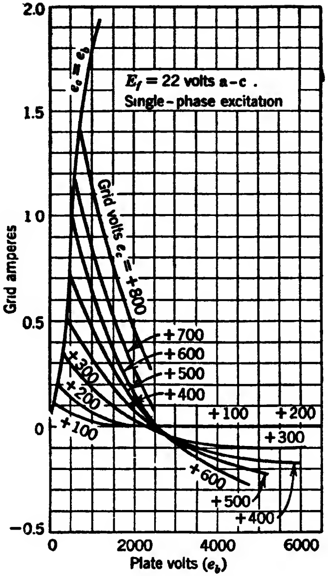
the volume of the tube increases more rapidly than the surface area. It is possible to obtain a greater output by operating glass tubes in parallel. A more direct solution is the use of water cooling or forced air cooling for tubes. In the water-cooled triode the plate is a round copper tube, which also constitutes the envelope of the tube as shown in Fig. 20. The copper plate is immersed in a bath of circulating water which conducts the heat away rapidly. It is sealed at the top to a glass envelope which serves to insulate and support the cathode and grid of the tube. This airtight, metal-to-glass union is called the housekeeper seal. In making this seal the open end of the copper plate is drawn out to a feather edge and placed in the molten glass of the envelope. Although there is some difference in the coefficient of expansion of the glass and copper, the seal along the thin part of the feather edge is seldom broken.

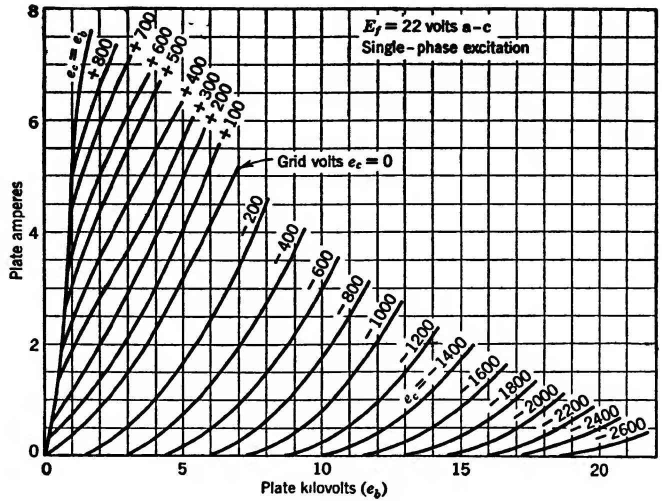
Water-cooled tubes are usually built in capacity ratings of from 1 to 100 kilowatts and plate-voltage ratings from 5000 to 100,000 volts. Tubes of 250-kilowatt rating have been built in the United States for use abroad. The rating and characteristics of a 10-kilowatt water-cooled triode are given in Fig. 20.

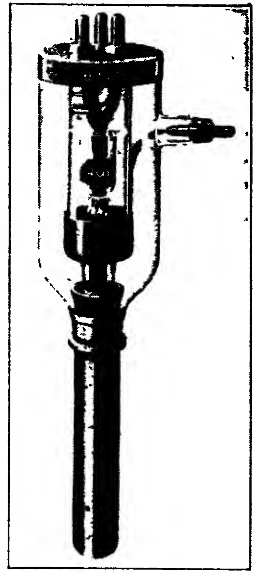
Water cooling of tubes involves some construction and maintenance problems, which has led to the development of power tubes using forced air cooling. A large capacity and moderately high-voltage tube of this type is illustrated in Fig. 21. The student should compare the construction details, rating, and characteristics of this tube with the preceding water-cooled unit.

The three-electrode vacuum tube may be classified as amplifier and relay, oscillator or alternating-current generator, modulator, and detector or demodulator. The particular function performed by the tube depends not upon its construction but upon the external circuits with which it is associated. Thus a single tube might be connected into circuits for serving as amplifier, oscillator, modulator, or detector. In practice certain types of tubes are chosen for the different applications, but this selection is not determined by any difference in their inherent theory of operation. The circuits and applications of triodes will be covered in succeeding chapters.

Limitations of a Triode. There are two properties of a triode and its circuit that limit its range of operation as an amplifier. The first of these arises from the capacitance coupling between the grid and the plate. The grid and the plate are two electrodes separated by an insulator (a vacuum), and thus they constitute a condenser. A con-







R-F POWER AMPLIFIER AND OSCILLATOR

General Data

Filament—two-section tungsten	
Voltage per section	11.0 volts
Current	60.0 amp
Amplification factor	8.0
Interelectrode capacitance	
Grid to plate	27.0 μμf
Grid to filament	18.0 μμf
Plate to filament	2 0 μμf
Mounting position—filament end up	

Class C Telegraphy

Maximum Ratings, Absolute Values

D-c plate voltage, max	12,000 volts
D-c grid voltage, max	−3,000 volts
D-c plate current, max	2.0 amp
D-c grid current, max	0.15 amp
Plate input, max	18 kw
Plate dissipation, max	6 kw

Typical Operation

D-c plate voltage	8,000	10,000 volts
D-c grid voltage	−1,800	−2,000 volts
Peak r-f grid voltage	2,500	2,900 volts
D-c plate current	1.1	1.45 amp
D-c grid current, approx	0.06	0.11 amp
Driving power, approx	150	310 watts
Power output, approx	6.5	10 kw

Fig. 20. Rating and characteristics of a water-cooled transmitting triode (Courtesy Radio Corporation of America.)

TRANSMITTING TRIODE

Forced Air Cooled

Electrical

Filament—three-section tungsten	
Excitation—1ϕ a-c, 3ϕ a-c, 6ϕ a-c, or d-c	
Voltage per strand	10 volts
Current per terminal	61 amp
Amplification factor	36
Direct interelectrode capacitances (approx)	
Grid to plate	34 $\mu\mu$f
Grid to filament	48 $\mu\mu$f
Plate to filament	3.5 $\mu\mu$f

Maximum Ratings, Absolute Values

D-c plate voltage, max	20,000 volts
D-c grid voltage, max	−3,000 volts
D-c plate current, max	4 amp
D-c grid current, max	0 4 amp
Plate input, max	70 kw
Plate dissipation, max	20 kw
Radiator temperature, max	180° C

Typical Operation

D-c plate voltage	12,000	15,000	18,000 volts
D-c grid voltage	−800	−900	−1,000 volts
Peak r-f grid voltage	1,430	1,520	1,630 volts
D-c plate current	3 5	3 6	3 6 amp
D-c grid current, approx	0 26	0 25	0.21 amp
Driving power, approx	360	370	340 watts
Power output, approx	30	40	50 kw

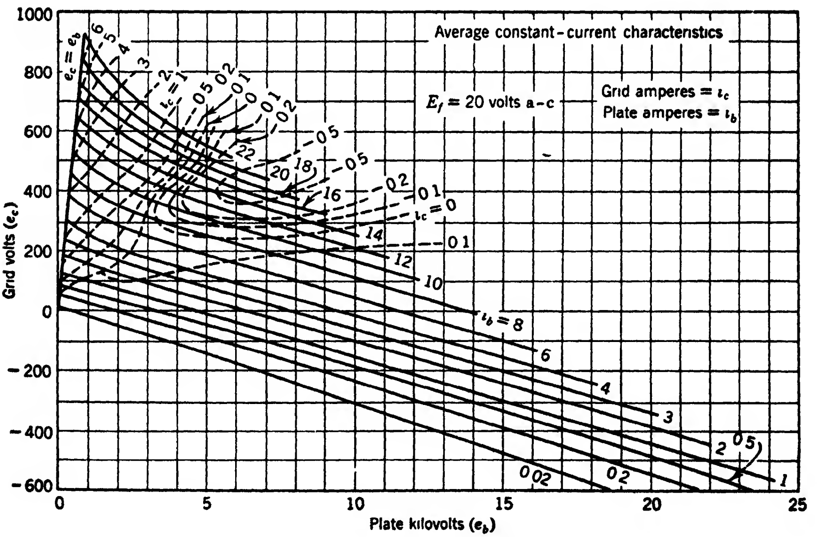


Fig. 21 Rating and characteristics of a power triode. (Courtesy Radio Corporation of America.)

denser is an open circuit to a direct current but a conductor for alternating currents. Thus, although the cathode-grid and cathode-plate circuits of Fig. 22 (left) appear to be separated, the interelectrode capacity between plate and grid forms a coupling between them. Through this coupling a feedback may occur from plate to grid caused by changes of potential (alternating) on the plate. This feedback will interfere with the normal control by the grid, or it may even start oscillations and cause the tube to serve as an oscillator.

The second limitation in the triode amplifier circuit results from

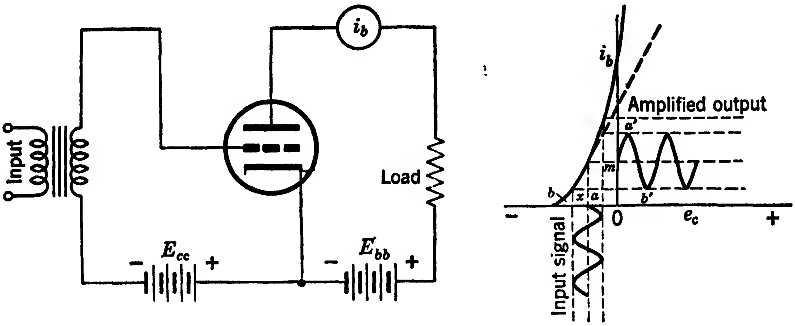


FIG. 22. Simple triode circuit and diagram illustrating the theory of amplification.

the changes in the cathode-plate potential with the changes in the plate current. In order to utilize the amplified signal, it is necessary to place a resistance, inductance, or transformer as a load in the plate circuit. (See Fig. 22.) This load has impedance and the plate current through this impedance produces a voltage drop ZI. A varying plate current produces a varying drop over the load impedance, which in turn is subtracted from the constant potential supply of the plate E_{bb}. This varying load drop occurs in all tubes, but its effect is especially undesirable in the triode. Here the changes in cathode-plate potential will vary the plate current as was pointed out in the discussion of the two-electrode tube (see Fig. 9, Chapter III, page 47). This change in plate current is a direct action on the plate-cathode circuit, whereas the interelectrode action suggested in the preceding paragraph takes place in the grid-cathode circuit.

The Tetrode. The limitations in the use of the triode as outlined in the preceding article can be overcome by the addition of a fourth electrode, a grid, placed between the regular or control grid and the plate. This fourth electrode is called the screen grid, and the tube is known as a screen-grid tube.

The construction of the screen-grid tube is shown schematically in Figs. 23 and 24. In the smaller receiving tube the screen grid consists of two grids connected at the top to form a nearly complete screen or shield for the plate. This construction reduces the interelectrode capacity between grid and plate to an average of $\frac{1}{800}$ of its value without the screen and overcomes the first limitation of the triode.

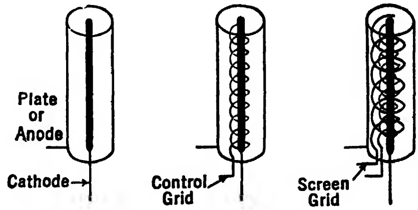


FIG. 23. Location of the screen grid in a vacuum tube.

A typical circuit for using a screen-grid tube is given in Fig. 24. The screen grid is maintained at a constant potential by connection to the cathode via a part of the plate supply voltage. Thus the screen grid sets up in the space surrounding it a constant potential. This potential supplies a constant accelerating force (attraction) for any electrons that pass the control grid. Normally nearly all electrons that reach the position of the screen grid pass through the meshes of the grid and land on the plate. *Moderate changes in the plate potential will not influence the number of electrons reaching the plate.* Thus the screen has served to overcome the second limitation of the triode. The screen grid attracts and gathers those electrons which are traveling directly toward its mesh wires, and these electrons constitute a small screen-grid current.

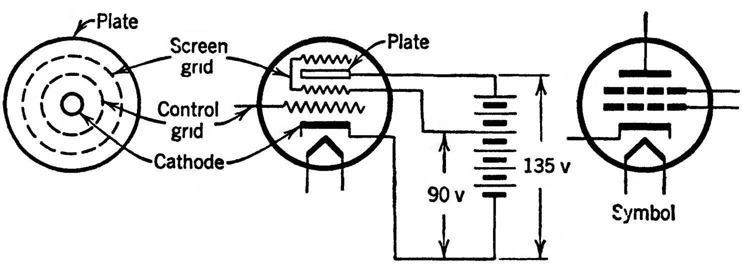


FIG. 24. Schematic diagram of a tetrode and its circuit.

The action of the screen grid may be pictured from the potential distribution curve of Fig. 25. The cathode is held constant at zero (ground potential) while the space surrounding the screen grid is held at another constant potential (say 90 volts). The control-grid potential is governed by the incoming signal and may swing from negative to positive. Likewise, the plate potential may swing from 90 to 135

volts. These latter swings will not disturb the constant potential at the screen grid and will not affect the electron flow to the plate. It should be noted that the actual potential distribution in a tetrode is not so simple as that indicated in Fig. 25. Space charges near the

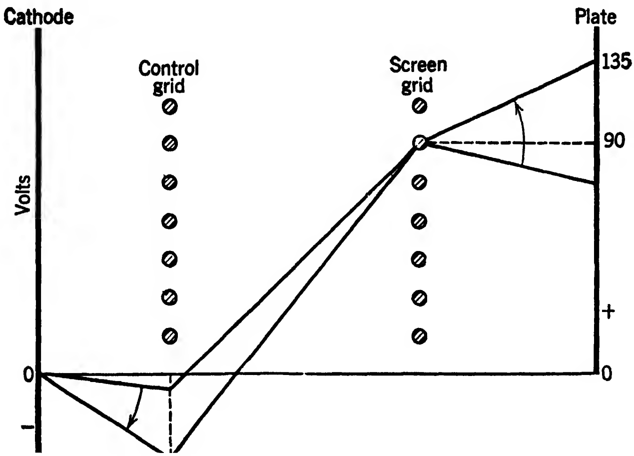


FIG. 25. Potential distribution in a screen-grid tube.

cathode or other electrodes, the geometry of the tube, and the intergrid wire regions will change the linear distribution which has been used to give a simple concept of the operation of the tube.

The screen-grid tube gives an unexpected performance when the plate potential falls to low values. This performance can be deter-

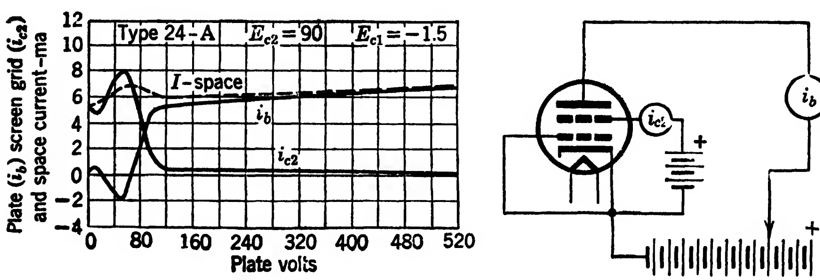


FIG. 26. Plate-current-voltage characteristic and circuit of a screen-grid tube.

mined by the circuit of Fig. 26 (right), from which the plate current can be determined for a variation of plate potential from zero to maximum rated value. The early types of screen-grid tubes (Type 24-A) gave the plate-current plate-voltage characteristic shown in

Fig. 26 (left), when the screen-grid voltage was held at normal value and the control grid kept at zero volts (cathode potential). At zero plate potential, the plate current is zero, as would be expected. Then, as plate potential rises to 10 or 15 volts, the plate current rises. With a further rise in plate potential, the current starts to drop and continues to do so until it becomes negative in value. After reaching a minimum point, the plate current rises with plate potential to a normal value and then becomes approximately constant for further rise in plate voltage. This unusual characteristic is due to secondary emission of electrons from the plate which are attracted to the screen grid. Secondary emission occurs regularly at the plate of all tubes owing to impinging electrons, but for the triode all the electrons of secondary emission are attracted back to the plate and hence do not produce any change in the current of the plate circuit. In the screen-grid tube, electrons are splashed out of the plate as shown in Fig. 27. If the screen grid happens to be at a higher potential than the plate and if the electrons are splashed out with sufficient velocity, some or all of these electrons of secondary emission will pass to the screen grid. If there are more electrons of secondary emission than original impinging electrons, the electron current may reverse in the plate circuit and flow to the screen grid. Thus the characteristic of Fig. 26 becomes easy to understand. Beginning at the left (zero plate potential), no electrons are attracted by plate. Then, as plate potential goes somewhat positive, electrons are attracted to it but without sufficient velocity to cause secondary emission. With a rise in plate potential a point is soon reached where secondary emission occurs with some electrons going to the screen grid. A further increase in plate potential increases the secondary emission rapidly and with it the number of electrons going to the screen grid. As the potential on the plate approaches that of the screen grid, the plate begins to receive back more electrons and a reversal of plate current trend takes place. When the plate potential passes the screen in magnitude, the plate current rises to full or normal value.

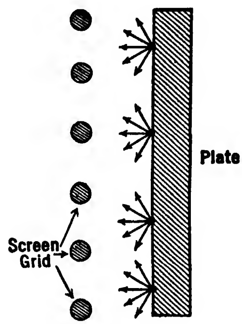


FIG. 27. Secondary emission in a tetrode.

The screen-grid current for the tube in question is shown by the i_{c2} curve of Fig. 26. This curve is the inverse of the plate current, as might be expected. The space current or sum of plate and screen-grid cur-

SCREEN-GRID VOLTAGE AMPLIFIER

Filament	
Voltage	2 0 volts
Current	1 6 amp
Amplification factor	85–170
Interelectrode capacitance	
Grid to plate	0 025 μμf
Input	4 5 μμf
Output	8 0 μμf
D-c plate voltage	135–180 volts
D-c screen voltage	45 67 5 volts
D-c grid bias	1 5–4 5 volts

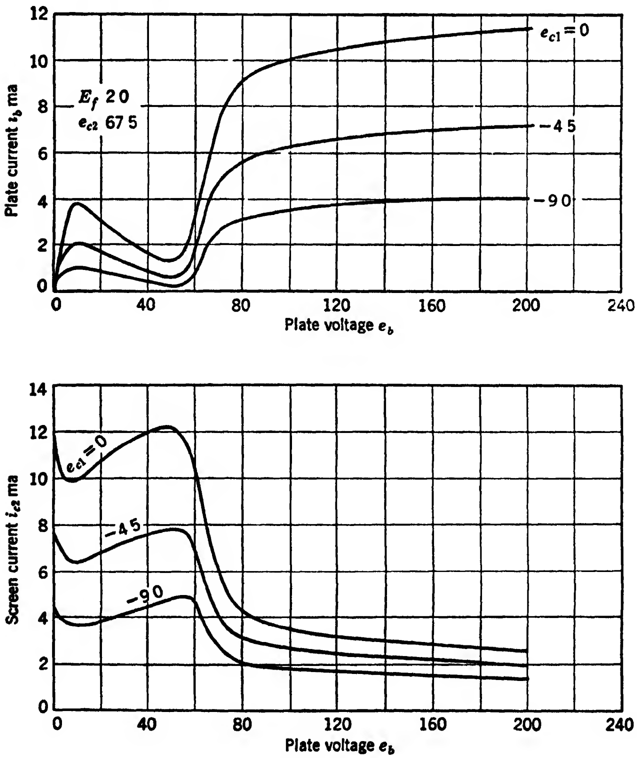


FIG. 28. Rating and characteristics of a voltage amplifier screen-grid tetrode. (Courtesy Western Electric Company.)

rent is approximately constant as shown by the curve of Fig. 26, because the electrostatic field produced by the screen grid is constant.

Screen-grid tubes of later manufacture have been improved by special treatment of the plate which reduces their tendency toward secondary emission. This improvement prevents the plate current from becoming negative and gives characteristic curves as shown in Fig. 28. It is evident from these characteristic curves that the screen-grid tube should give more linear amplification as long as the plate potential exceeds the screen-grid potential, and that wide swings in plate potential may occur in that region. For lower values of plate potential the tube will give unsatisfactory service. The effect of secondary emission in the screen-grid tube can be overcome in two different ways, to be covered in succeeding articles. The principal application of screen-grid tubes is for radio-frequency amplifiers. The rating and characteristics of a typical screen-grid tube are given in Fig. 28.

The Pentode. The pentode is a tube having five electrodes. The fifth electrode, called the suppressor grid, was added to overcome the

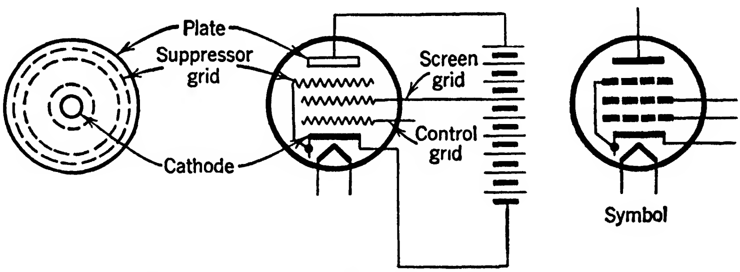


FIG. 29. Schematic diagram and circuit of a pentode.

harmful effect of the secondary emission in the screen-grid tube. The suppressor grid is a mesh placed between the plate and screen grid as shown in Fig. 29. It is usually connected directly to the cathode inside the tube, though it may be brought outside for other connections. When connected to the cathode the suppressor grid serves as a shield or suppressor to prevent the electrons of secondary emission from passing to the screen grid. The suppressor grid performs this function by creating a field of near zero potential through which electrons (secondary emission) must pass before coming under the influence of screen grid. This explanation may become more clear from a study of Fig. 30, which gives a schematic potential distribution for a pentode. Here point P represents the constant zero potential on the suppressor grid. Electrons splashed out of the plate must move away from it

against the attraction of its positive potential and past point P before they come under the influence of the potential on the screen grid.

The plate-current plate-potential characteristics of pentodes are given in Figs. 31 and 32. The pentode permits wide swings of plate potential without affecting the fidelity of amplification. The action of the suppressor permits large power output for low input voltage on the control grid and it permits high-voltage amplification at moderate values of plate voltage. Accordingly, some pentodes have a large

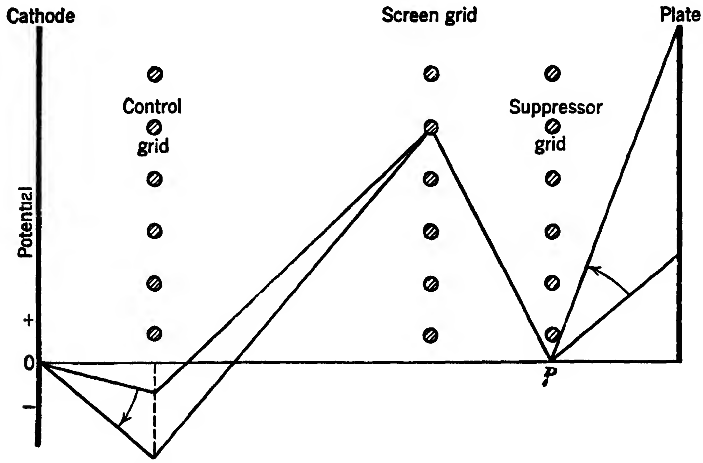
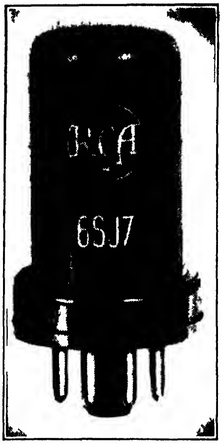


FIG. 30. Potential distribution in a pentode.

power output and are used in the final stage of audio amplification for supplying the current to the loudspeaker. Other types of pentodes are used as voltage amplifiers. The construction of typical pentodes is illustrated in Figs. 31, 32, and 33. The rating and characteristics of two pentodes are given in Figs. 31 and 32.

In both pentodes and tetrodes, the control grid is sometimes designed to give a variable mu. This is secured by a close spacing of the grid wires at the top and bottom of the grid and a coarse spacing at the center. The effect of this construction upon the plate-current characteristic is illustrated in Fig. 32 (upper right). The characteristic curves of the tube near the cutoff point are broadened and the tube can be used for amplification at lower levels of plate current. For comparison purposes a curve for a normal grid is shown by the dotted line marked X. Tubes employing the variable-mu type of grid are used for grid-bias volume control in radio-frequency amplifiers.

TRIPLE-GRID DETECTOR-AMPLIFIER



Heater—coated		
Voltage (a-c or d-c)		6 3 volts
Current		0 3 amp
Direct interelectrode capacitance		
Pentode Conn	Grid to plate, max	0 005 $\mu\mu$f
	Input, max	6 0 $\mu\mu$f
	Output, max	7 0 $\mu\mu$f
Triode Conn	Grid to plate, max	2 8 $\mu\mu$f
	Grid to cathode, max	3 4 $\mu\mu$f
	Plate to cathode, max	11 $\mu\mu$f

Amplifier (Pentode Connection)

Plate voltage, max	300 volts
Screen voltage, max	125 volts
Screen supply voltage, max	300 volts
Grid voltage, min	0 volts
Plate dissipation, max	2 5 watts
Screen dissipation, max	0 3 watt

Typical Operation and Characteristics Class A_1 Amplifier

Plate	100	250 volts
Screen	100	100 volts
Grid	−3	−3 volts
Suppressor—connected to cathode at socket		
Plate resistance	0 7	meg
Transconductance	1575	1650 μmhos
Grid bias (for plate current = 10 μamp)	−8	−8 volts
Plate current	2 9	3 ma
Screen current	0 9	0 8 ma

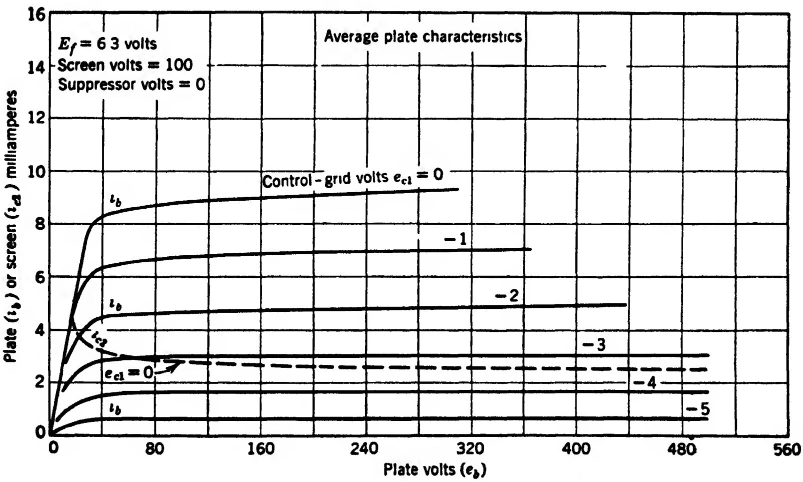


Fig 31 Rating and characteristics of a pentode detector-amplifier (Courtesy Radio Corporation of America)

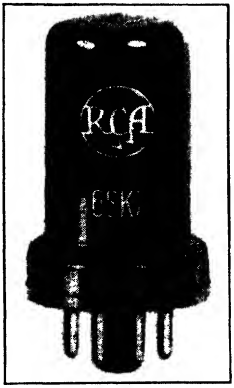
TRIPLE-GRID SUPERCONTROL PENTODE

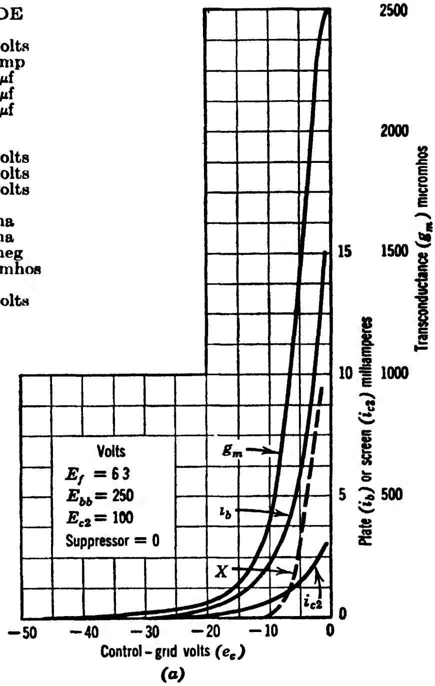
Heater voltage (a-c or d-c)		6.3 volts
Heater current		0.3 amp
Grid-plate capacitance,* max		0.003 $\mu\mu f$
Input capacitance *		6 $\mu\mu f$
Output capacitance *		7 $\mu\mu f$

Typical Operation

Plate voltage	100	250 volts
Screen voltage	100	100 volts
Grid voltage	−3	−3 volts
Suppressor—connected to cathode at socket		
Plate current	8.9	9.2 ma
Screen current	2.6	2.4 ma
Plate resistance, approx	0.25	0.8 meg
Transconductance	1900	2000 μmhos
Grid bias for transconductance of 10 μmhos	−35	−35 volts

* With shell connected to cathode.





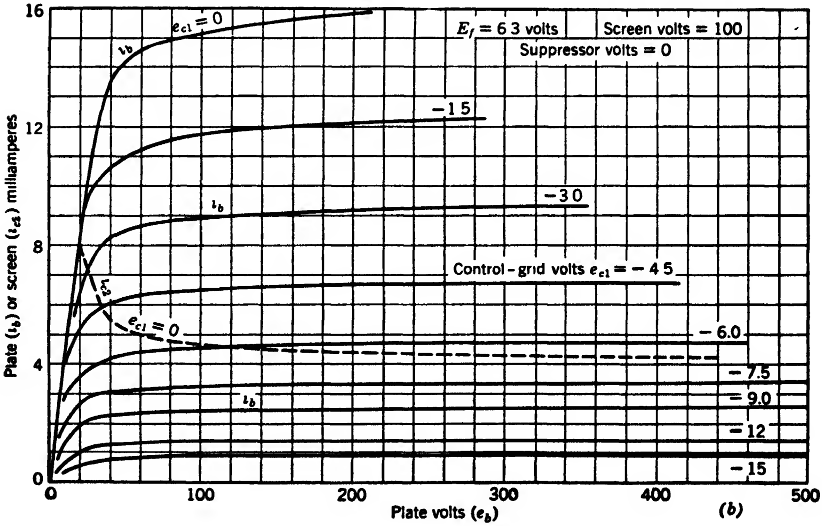


FIG. 32. Rating and characteristics of a triple-grid super-control pentode. (Courtesy Radio Corporation of America.)

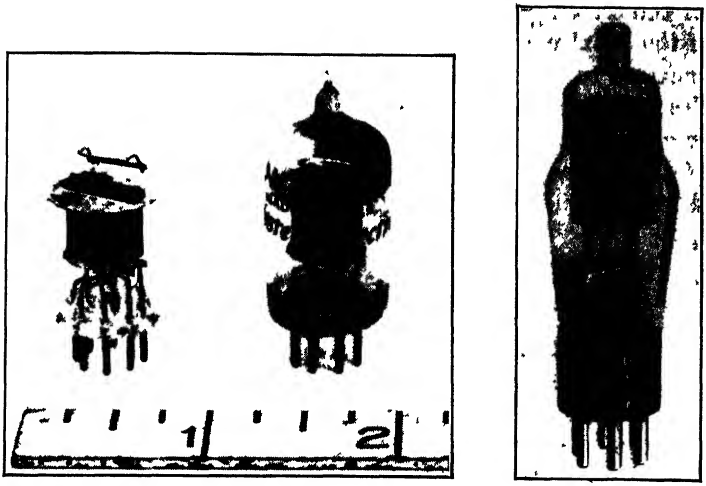


FIG. 33. Pentodes. Those on the left are for low-power applications at high and ultra-high frequencies. (Courtesy Western Electric Company.)

Beam Power Tube. A beam power tube is a tetrode * in which directed electron beams increase the power capacity and operating characteristics of the tube. This tube has three special features in its

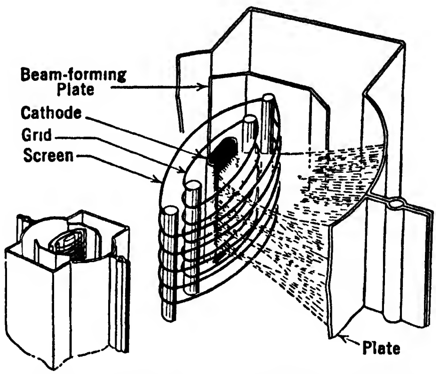


FIG. 34. Construction of a beam power tube.

construction. First, the screen and control grid are composed of wires wound in helices so that each turn of the screen is shaded from the cathode by a turn of the control grid. This careful alignment tends to pass the electrons to the plate in beams and serves to reduce the magni-

* If beam-forming plates are called electrodes, the tube may be considered a pentode.

tude of the screen-grid current. The second feature in the construction of the beam power tube is the use of beam-forming plates which are connected to the cathode (Fig. 34). These plates serve to prevent any electrons from leaving the grid near its end supports and serve to give a sharp cutoff for the beams. The third feature of construction of the beam power tube is that the screen grid and plate are spaced relatively far apart (Fig. 35) and that these electrodes are

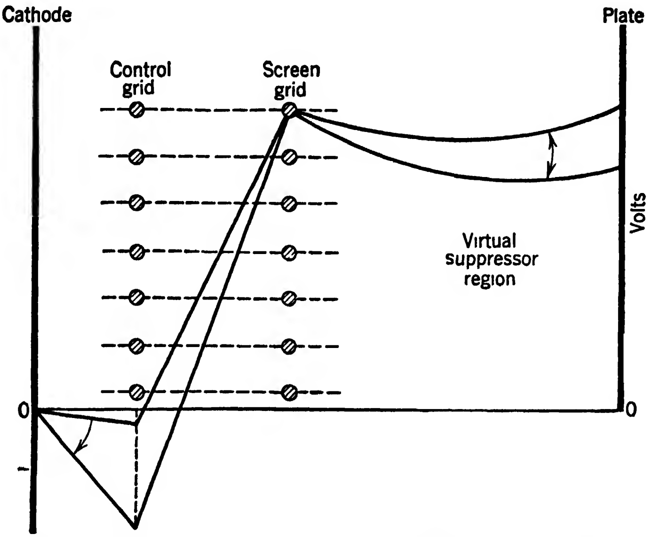
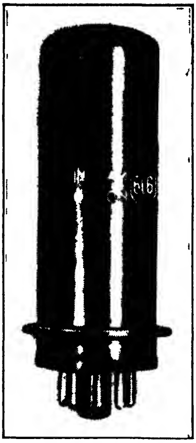


FIG. 35. Potential distribution in a beam power tube.

operated at approximately the same potential. This construction results in a suppressor action between the screen and plate. To understand this action, assume that screen and plate are at the same potential and that the cathode is cold (zero emission). Under this condition the potential distribution between the screen and plate will be uniform and no potential gradient will exist. Next assume similar conditions with normal emission from the cathode and a control-grid potential that permits a normal plate current. Now moving electrons are present everywhere between screen and plate, and these electrons constitute a negative space charge which is strengthened by the action of the suppressor plates and by the fact that the electrons exist in concentrated beams. This negative space charge lowers the potential in the region between screen and plate and results in *a change of pace of the moving electrons.* This change of pace results in a variation of density of the electrons in the screen-plate space with a corresponding



BEAM POWER AMPLIFIER

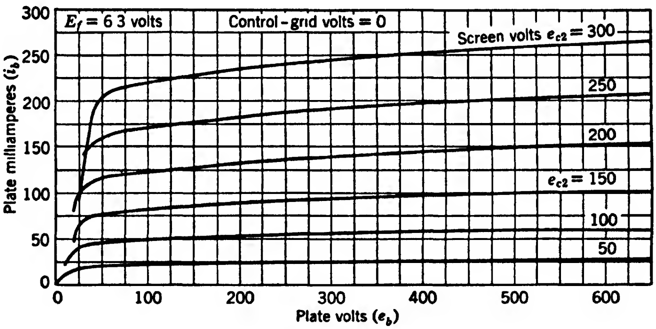
Heater—coated unipotential cathode		
Voltage (a-c or d-c)	6.3 volts	
Current	0.9 amp	

SINGLE-TUBE AMPLIFIER—CLASS A_1

Plate voltage, max	360 volts
Screen voltage, max	270 volts
Plate dissipation, max	19 watts
Screen dissipation, max	2.5 watts

Typical Operation

	Fixed Bias		Cathode Bias		
Plate	250	350	250	300	volts
Screen	250	250	250	200	volts
Grid	−14	−18		. . .	volts
Cathode resistor			170	220	ohms
Peak a-f grid voltage	14	18	14	12.5	volts
Zero signal plate current	72	54	75	51	ma
Max signal plate current	79	66	78	54.5	ma
Zero signal screen current	5	2.5	5.4	3	ma
Max signal screen current	7.3	7	7.2	4.6	ma
Plate resistance	22,500	33,000	.	.	ohms
Transconductance	6,000	5,200	. .	. .	µmhos
Load resistance	2,500	4,200	2,500	4,500	ohms
Total harmonic dist.	10	15	10	11	%
Max signal power output	6.5	10.8	6.5	6.5	watts



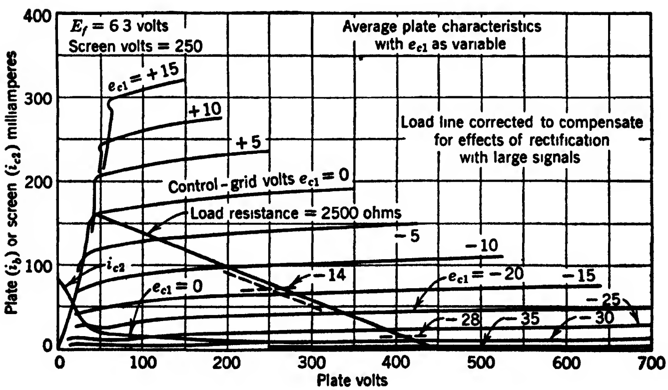


FIG. 36. Rating and characteristics of a beam power amplifier. (Courtesy Radio Corporation of America.)

change of potential distribution. The net result of this action is the formation of a virtual suppressor between the screen and plate as shown in Fig. 35.

The rating and characteristics of a typical beam power tube are given in Fig. 36. It should be noted that wide changes of plate potential above 75 volts have little effect upon plate current. Also, the lower set of curves shows a sharp bend at the knee of plate-current plate-voltage curve. This means that little harmonic distortion occurs above the bend.

The advantages of tubes of the beam power type are high power output, high power sensitivity, and high efficiency. They are frequently used in the output stage of radio receivers and for other forms of load.

Pentagrid Converter. The pentagrid converter is a multielectrode tube having one cathode, one plate, and five grids. It is a dual-purpose tube which does not involve any new theory of action above that covered in preceding types. It is commonly used as a frequency converter, and the use of five grids gives rise to the name pentagrid converter. Frequency conversion involves the action of an oscillator and a modulator. The circuits for the pentagrid converter combine these two functions within a single tube, and the electron stream serves as the coupling.

Cathode-Ray Tube. A cathode-ray tube is a device for producing electron beams and for projecting them upon a fluorescent screen to give a picture of some electrical phenomenon. The first device of this type was developed by Braun in 1897. Cathode-ray tubes have used (1) cold electrodes with high potentials in vacuum, (2) hot cathodes with low gas pressure and low voltage between electrodes, and (3) hot cathodes in vacuum with fairly high accelerating potentials. Nearly all the tubes in use today fall in the third class.

A typical cathode-ray tube with electrostatic controls embodies the schematic construction shown in Fig. 37. The seven electrodes in this device serve to produce electrons, to concentrate them in a small beam, and to aim that beam upon various parts of the fluorescent screen on the end of the tube. The first five electrodes on the left constitute what is called the electron gun. The first electrode on the left is the cathode, a barium-coated plate with a heater behind it. Electrons from the cathode pass through a small hole in the control grid. Next, these electrons are accelerated by a higher potential on the second grid which is a disk with a small hole at its center. The first anode is a tube with two baffles, each containing a tiny hole at the center. This first anode has a potential of several hundred volts and serves to

accelerate and concentrate the beam of electrons. The second anode has a still higher potential, usually in thousands of volts, and constitutes the final accelerating stage in the electron gun. After emerging from the second anode, the electron stream passes between the two pairs of deflecting plates. A difference of potential on one pair of plates will produce an electric field which will exert a force upon the electrons in the stream and deflect them in their path. Similarly, a difference of potential on the other pair of plates placed at right angles to the first will likewise deflect the electron stream. Thus through the medium of these two pairs of deflecting plates the electron beam can be

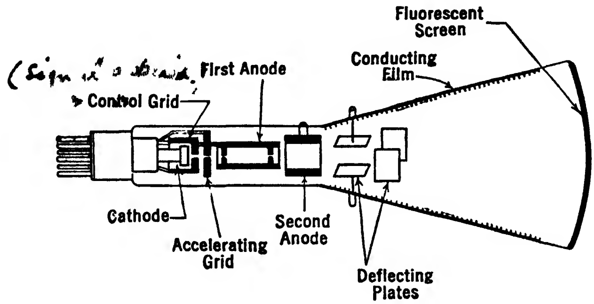


FIG. 37. Construction of a cathode-ray tube having electrostatic controls.

directed upon any part of the fluorescent screen. The trigger for the electron gun is the control grid which receives an electrical input signal. The varying potential on this grid can start, stop, and control the intensity of the beam at all times. The electron beam impinging on the screen causes it to fluoresce in proportion to the intensity of the beam. Due to the persistence of vision of the eye, a moving beam produces a trace of light or, in the case of television, a complete image on the fluorescent screen. The fluorescent screen consists of a coating of a phosphor. Phosphors consist of compounds such as zinc silicate, cadmium tungstate, zinc sulphide, cadmium sulphide, and calcium tungstate, or a mixture of these compounds together with some substance such as silver or copper. The choice of the phosphor is determined by the color and the persistence time desired. A conducting coating on the inside of the envelope of the tube serves to return the electrons from the fluorescent screen to one of the anodes. The secondary emission due to the impinging electrons removes electrons from the screen and prevents the building up of a negative charge.

Many cathode-ray tubes use magnetic fields for focusing and deflecting the electron beam. The construction of a tube of this type is shown

schematically in Fig. 38. The electron beam is produced and accelerated by means of an electrostatic electron gun as explained in the preceding paragraph. The magnetic focusing is produced by a solenoid surrounding the tube. The coils for producing the fields for magnetic deflection are placed on the inside of a hollow cylinder like the stator coils in a-c machines. The magnetic focusing and deflection coils are usually formed into a single hollow cylindrical unit which slips over

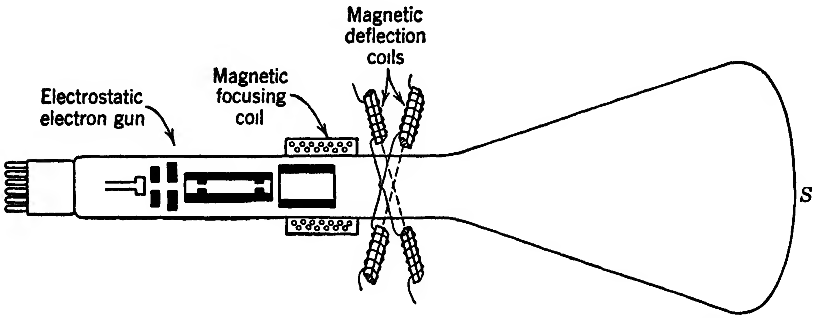


FIG. 38. Construction of a cathode-ray tube using magnetic deflection and focusing.

the cathode-ray tube. This unit is called a yoke. The general theory of electrostatic and magnetic focusing and deflection employed in cathode-ray tubes was covered under the subject of electron ballistics on pages 7–17.

The principal applications of the cathode-ray tube are in oscilloscopes and in television receivers. In the oscilloscope a trace or curve is made on the screen which depicts the changes of voltage or current with time or some other variable. These traces (Fig. 39) can be viewed with the eye or photographed. In the television receiver a rapid scanning of the screen gives the impression of a complete image on the screen. The cathode-ray tube is very important as a measuring instrument for high-frequency changes. It holds this position because a beam of electrons or cathode rays passing from the cathode to the target possesses so little inertia that the beam can be deflected at a rapid rate, thus making possible the observance of high-frequency and transient phenomena. The power required to deflect the beam is negligible, and hence studies of weak signals can be made because so little power is required and the circuit conditions are not affected by the measuring equipment.

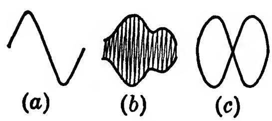
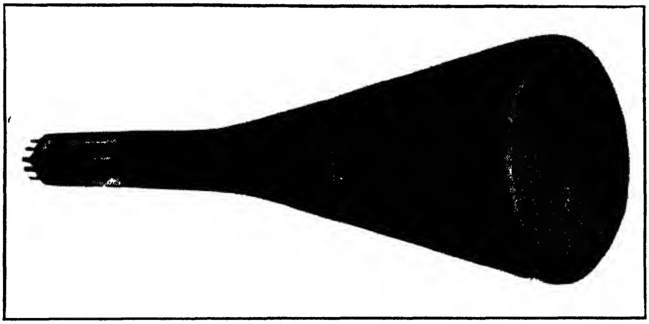


FIG. 39. Traces shown by an oscilloscope.



CATHODE-RAY TUBE

Heater, for unipotential cathode		
Voltage (a-c or d-c)		6.3 ± 10% volts
Current		0.6 amp
Direct interelectrode capacitances (approx)		
Grid No. 1 to all other electrodes		8.0 μμf
DJ1 to DJ2		1.3 μμf
DJ3 to DJ4		1.2 μμf
DJ1 to all other electrodes		9.5 μμf
DJ3 to all other electrodes		12.0 μμf
DJ1 to all other electrodes except DJ2		8.0 μμf
DJ2 to all other electrodes except DJ1		7.5 μμf
DJ3 to all other electrodes except DJ4		10.0 μμf
DJ4 to all other electrodes except DJ3		7.5 μμf
Phosphor		No. 1
Fluorescence		Green
Persistence		Medium
Focusing method		Electrostatic
Deflection method		Electrostatic
Maximum Ratings, Absolute Values		
Anode No. 2 and grid No. 2 voltage, max		2200 volts
Anode No. 1 voltage, max		1100 volts
Grid No. 1 (control electrode) voltage		
Negative value, max		125 volts
Positive value, max		0 volts
Peak voltage between anode No. 2 and any deflecting electrode, max		550 volts
Typical Operation		
Anode No. 2 and grid No. 2 voltage	1500	2000 volts
Anode No. 1 voltage for focus at 75% of grid No. 1 voltage for cutoff	337	450 volts
Grid No. 1 voltage for visual cutoff	−30	−40 volts
Deflection sensitivity		
DJ1 and DJ2	0.404	0.303 mm/d-c volts
DJ3 and DJ4	0.446	0.334 mm/d-c volts
Deflection factor		
DJ1 and DJ2	63	84 d-c volts/in
DJ3 and DJ4	57	76 d-c volts/in

FIG. 40. Cathode-ray tube. (Courtesy Radio Corporation of America.)

Average characteristics

$E_f = 6.3$ volts
Anode No. 1 volts adjusted to give focus

Curve	Electrode current	Anode No. 2 and grid No. 2 volts
A	Anode No. 1	2000
B	Anode No. 1	1500
C	Anode No. 2 and grid No. 2	2000
D	Anode No. 2 and grid No. 2	1500

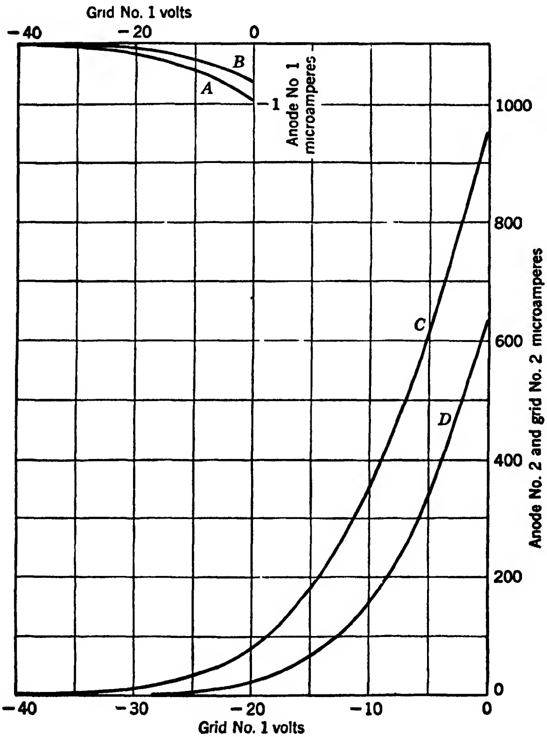


FIG. 40 (*continued*). Cathode-ray tube characteristics. (Courtesy Radio Corporation of America.)

Nearly all cathode-ray tubes in use today are of the high-vacuum type. The functioning of these tubes is controlled by electrostatic and magnetic fields. Cathode-ray tubes are special tubes, but their special importance in the oscilloscope, which is used so widely in electrical measurements and tests, seemed to justify the introduction of their construction at this point. The rating, characteristics, and picture of a typical electrostatic cathode-ray tube are given in Fig. 40.

Electron-Ray Tube. The electron-ray tube is a combination of a triode and simple cathode-ray tube built into one envelope. Sectional

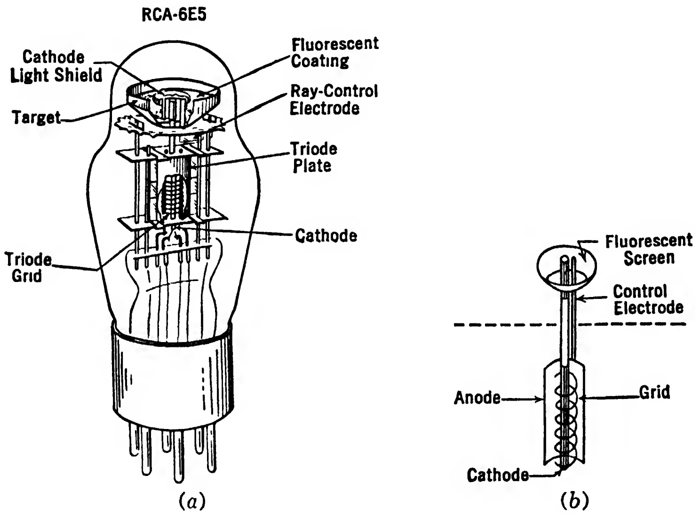


FIG. 41. Construction of an electron-ray tube. (Courtesy Radio Corporation of America.)

views of this device are given in parts *a* and *b* of Fig. 41. The upper or dome-shaped part of the tube contains the cathode-ray part of the device, and the lower part houses the triode. A common heater-type cathode passes through the vertical axis of both sections, but the oxide coating covers the parts shown shaded only. In the upper section, the anode or target consists of a circular plate having a fluorescent coating on the upper side and a hole in its center. When a positive potential exists on this plate, electrons from the cathode are attracted to it and cause it to fluoresce with a green color. Without any other influence the circular plate would show a green circle of light with a black center. A grid called a ray-control electrode passes between the cathode and the inner circle of the fluorescent disk. This grid consists of a thin strip

of metal placed parallel to the vertical cathode but insulated from both the cathode and the fluorescent disk. Now if this grid has the same potential as the disk, it will have little influence on the electrons going to the disk and a circle of green light still exists. However, if the grid becomes negative with respect to the disk, it will repel the electrons leaving the cathode and will cast a shadow on the circle as shown in Fig. 42. It should be noted that the grid or ray-control electrode for the top section is really an extension of the plate of the triode in the lower section. Hence the potential on the triode plate controls the potential of the grid of the top section.

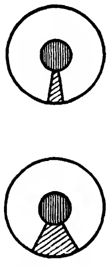


FIG. 42. Indications given by electron-ray tube.

The electron-ray tube sometimes carries the commercial name of the "magic eye" because of the resemblance to an eye suggested in Fig. 42. It is used as an indicator in sensitive electrical measuring devices and circuits, and in radio receivers to indicate the sharpness of the tuning for an incoming carrier signal. The sharpest tuning occurs when the shadow angle is at a minimum.

Mechano-Electronic Transducer. A new type of vacuum triode (5734) designed to provide a method of translating mechanical vibrations into electrical current variations was introduced by the Radio Corporation of America in 1948. This device, known as a *mechano-electronic transducer,* weighs $\frac{1}{16}$ ounce and is $1\frac{9}{16}$ inches long and $\frac{5}{16}$ inch in diameter.

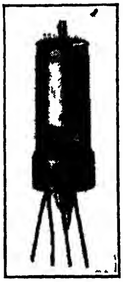
The construction of the mechano-electronic transducer is shown in the upper right-hand corner of Fig. 43. A hollow metal envelope is sealed at one end with a flexible metal diaphragm and with a glass press at the other end. The plate consists of a cone-shaped shaft sealed into the metal diaphragm to permit mechanical movement. An oxide-coated heater type of cathode lies parallel to the plate. A grid is spaced between the cathode and plate as shown.

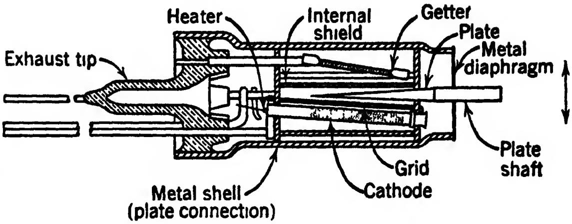
The device is operated by a suitable mechanical connection to the plate shaft by means of which any rocking movement or vibration will vary the geometry of the tube, and thus change the magnitude of the cathode-plate current. Thus the plate current is varied mechanically instead of by electrostatic control as in the conventional triode. A typical cathode-plate circuit employs a plate supply of 300 volts and a load resistor of approximately 75,000 ohms across which the output voltage is developed. The grid may be connected directly to the cathode, or a small negative bias up to 2 volts may be

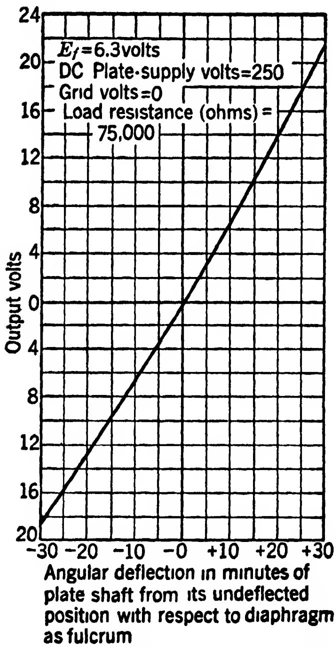
MECHANO-ELECTRONIC TRANSDUCER

TRIODE TYPE

Tentative Data







GENERAL DATA

Electrical	
Heater, for unipotential cathode	
Voltage (a-c or d-c)	6.3 volts
Current	0.15 amp
Mechanical	
Mounting position	Any
Deflection of plate, max	±0.5 deg
Overall length, max	1 300 inches
Diameter, max	0.328 inches
Maximum Ratings, Design-Center Values	
D-c plate supply, max	300 volts
D-c plate current, max	5 ma
Plate dissipation, max	0 4 watt
Peak heater—cathode voltage	
Heater ± with respect to cathode, max	90 volts
Typical Operation	
D-c plate-supply voltage	300 volts
D-c grid voltage	0 volts
Amplification factor	20
Plate resistance	72,000 ohms
Transconductance	275 μmhos
D-c plate current	1.5 ma
Load resistance	75,000 ohms
Deflection sensitivity	40 volts/deg

FIG. 43. Rating and characteristics of mechano-electronic transducer. (Courtesy Radio Corporation of America.)

employed to exercise some control of the characteristics or to control the calibration of any associated circuits.

The small amount of inertia of the plate permits the measurement of vibrations up to 12,000 cycles per second. A maximum permissible plate deflection of 1 degree (½ degree from the undeflected position) will produce a variation of 40 volts across the load resistor. The output voltage characteristic and the general data for the mechano-electronic transducer are given in Fig. 43.

The mechano-electronic transducer will be important in electronic and control applications. At the time these lines are written many possible uses are being explored. Following are some possible applications: (1) measurements of minute deflection of beams and other structural shapes; (2) sorting of mechanical products that must be held to close dimensions; (3) vibration measurements; and for a suitable modification of the device described above (4) a pick-up of sound from phonograph disk records or direct from sound in air.

Problems

For the solution of problems requiring the use of characteristic curves of tubes, it is suggested that the curve sheet be covered with transparent paper held by short strips of masking tape. Light construction lines on this cover paper will not mar the original curves, and the overlay may be removed to serve as a part of the solution of the problem.

1. Determine the amplification factor of the triode in Fig. 11 for i_b values of 4 and 6 milliamperes, using the transfer characteristic.

2. Determine the amplification factor of the triode in Fig. 11 for i_b values of 8 and 10 milliamperes, using the average plate characteristic curves.

3. Determine g_m for the triode in Fig. 11 for $e_b = 200$ volts and i_b near 6 milliamperes; also for $e_b = 300$ with i_b near 3 milliamperes.

4. Calculate r_p for the triode in Fig. 11 for values of e_c at 0, -10, and -24 volts.

5. Calculate the amplification factor of the triode in Fig. 20 for e_c values near -1000, -1600, and -2000 volts.

6. Determine g_m for the triode of Fig. 20 at $e_p = 10{,}000$ volts.

7. Calculate r_p for the triode of Fig. 20 for $e_c = -1000$ and $e_b = 12{,}000$ volts.

8. Calculate r_p for the tetrode of Fig. 28 for $e_{c1} = 0$ and plate voltages of 20, 67, and 160 volts. Explain answer.

9. A tube has an amplification factor μ of 8 and an r_p of 9500 ohms for a certain condition of operation. What is g_m in micromhos?

10. From the following data find approximate values of μ, r_p, and g_m at the point $e_b = 180$ volts, $e_c = -12.5$ volts.

e_b (volts)	ϵ_c (volts)	i_b(ma)
180	−12.5	7.5
160	−10.0	7.5
180	−12.3	7.84

11. The pentode of Fig. 31 has a plate resistance of 7×10^5 ohms and a transconductance of 1575 micromhos. What is the value of mu for this condition of operation?

12. Anode No. 2 of the cathode-ray tube is operated at a potential 2000 volts higher than the cathode. With zero potentials applied to the deflection plates, what is the velocity of the electrons when they hit the fluorescent screen? If the deflection plates are parallel and extend 2 centimeters along the axis of the electron beam, what angle of deflection will be given to the beam by a uniform field of 50 volts per centimeter?

13. Assume that magnetic deflection coils are substituted in the preceding problem, producing a uniform field of 4.5 gauss for a distance of 4 centimeters along the axis of the tube. Calculate the angle of deflection for the electron beam after it has passed through this field (one pair of coils energized).

Chapter V

LINEAR AND NONLINEAR CHARACTERISTICS

The performance of electronic equipment is determined by the characteristics of the associated circuits as well as by the characteristics of electron tubes. An understanding of communication and industrial control by means of electronic devices requires a thorough knowledge of electric circuit theory. The reader is presumed to have such knowledge, and hence only a short summary of a-c theory plus a discussion of some phenomena of special importance in electronic circuits will be reviewed at this point.

A-C Theory. An a-c circuit may consist of resistors, inductors, capacitors, or combinations of these three units. The characteristics of these units are reviewed in Fig. 1. For a non-inductive resistance the current time change is in phase with the voltage; in a pure inductance the current lags the voltage by 90 electrical degrees; in a capacitance the current leads the applied voltage by 90 electrical degrees.

Circuits and circuit elements may be classified as *linear* and *nonlinear*. Under common usage a circuit or element is considered linear when the voltage-current relation follows Ohm's law $(I = E/Z)$. Thus a pure resistance, pure inductance, and a capacitance are generally considered linear elements in an a-c circuit. Similarly electronic devices that have unilateral conductivity, changes in amplifying power, changes in mutual conductance, and changes in plate resistance are usually classed as nonlinear elements. Under operating conditions the magnitude of circuit elements often becomes a function of temperature, frequency, flux density, and other factors. Accordingly, an analysis of the properties of circuit elements under varying conditions becomes desirable in the study of electronic control and communication circuits.

Nonlinear Characteristics of Resistors. The resistance of pure metals rises with temperature and for the usual operating range may be considered a straight-line function. Certain alloys of metals and some combinations of materials have variations of resistance with temperature which make them useful in control circuits. Some al-

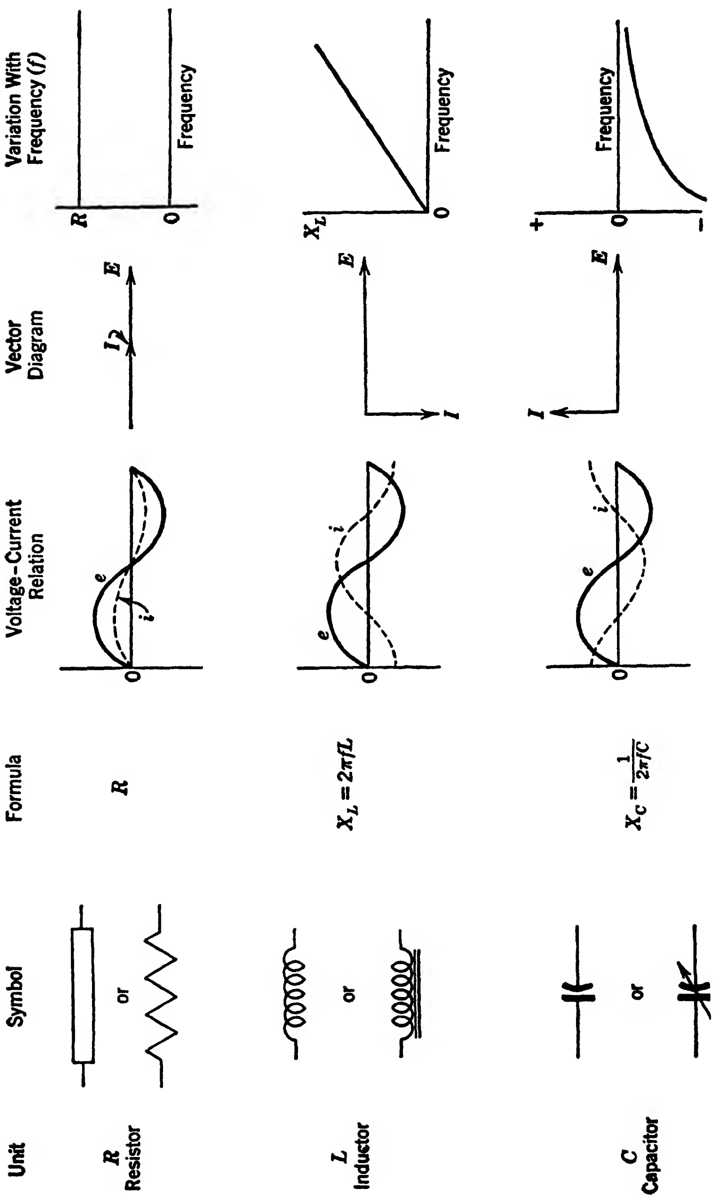

Unit
Symbol
Formula
Voltage-Current Relation
Vector Diagram
Variation With Frequency (f)
R
Resistor
or
R
e
i
0
I
E
R
0
Frequency
L
Inductor
or
$X_L = 2\pi fL$
e
i
0
E
I
X_L
0
Frequency
C
Capacitor
or
$X_C = \frac{1}{2\pi fC}$
e
i
0
E
I
+
0
-
Frequency

FIG. 1

loys have a zero temperature coefficient, whereas others have a negative temperature coefficient. Some rectifiers, such as those of crystal, copper oxide, and selenium, have a unilateral characteristic (see Chapter XIII), and others like thyrite and thermistors have negative resistance characteristics with the rise of voltage or current. For the low frequency used in power circuits the magnitude of resistors may be safely considered constant as shown in the upper right corner of Fig. 1. However, for the higher frequencies to be found in electronic circuits the ohmic resistance of wire-wound resistors rises with frequency due to the so-called skin effect. This effect is illustrated for a single straight conductor in the five views of Fig. 2. In part 1 of Fig. 2, it is

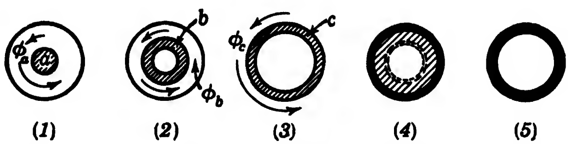


FIG. 2. Skin effect due to a-c current in straight conductors.

assumed that an instantaneous current of uniform density is flowing throughout the cross section of the conductor. This current will produce a magnetic field around the elemental cylinder a as shown, and the changing flux will induce an emf which will oppose the flow of current in a. In part 2 the elemental tube b carries current which causes flux ϕ_b to surround it. The outer tubular element c is surrounded by the external flux ϕ_c shown in part 3. Cylinder a is surrounded by the flux produced by itself plus that produced by the current flowing in the tubes b and c. Hence in part 1 the flow of current is opposed by the maximum or total emf of self-inductance due to the total surrounding flux. The current in b is opposed by the flux surrounding it. The effect of this self-inductance within the conductor is to distort the current density within the conductor, crowding the current toward the outside of the conductor as shown in parts 4 and 5 of Fig. 2. For very high frequencies nearly all the current flows on the surface so that a thin copper tube may have nearly the same resistance as a solid copper conductor of the same diameter. A thin coating of graphite or other conducting material on the surface of a cylinder of insulating material will make a suitable resistor or conductor for high frequencies. When a conductor is wound into a solenoid containing an air or magnetic core, the resulting flux will produce more or less distortion of the a-c current flowing in the conductors. Such distortion will add to the effective resistance of the

coil. It may be impossible to calculate the effective resistance of resistors at high frequencies, but such resistance may be determined under any given set of conditions through the use of the equation.

$$\text{Watts} = I^2 R_{\text{effective}}$$

$$R_{\text{effective}} = \frac{\text{watts}}{I^2} \tag{1}$$

Nonlinear Reactances. The magnitude of an inductive reactance $(2\pi fL)$ of a pure inductance L varies with frequency linearly as de-

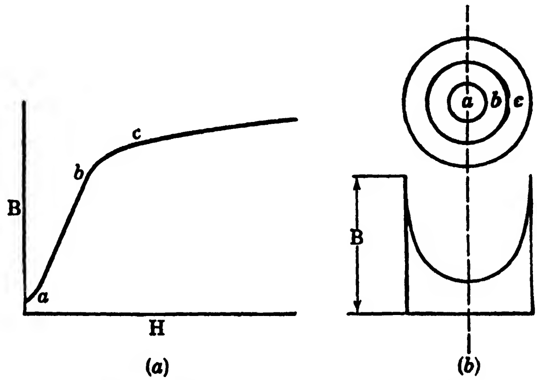


FIG. 3. Magnetic skin effect in circular iron core.

picted in Fig. 1. It is impossible to construct an inductance or choke coil which does not have some ohmic resistance because the coil consists of turns of metallic conductors which have resistance. Thus the impedance of inductance coils may be resolved into a component of resistance and a component of reactance in series. For most applications in electronics it is desirable to keep the resistance of choke coils low and the inductance L high, making the ratio L/R important in coil design. *The ratio $2\pi fL/R$ is called the figure of merit, or Q, of a coil.* In general, it is desirable to use coils having a Q of 10 or more for audio frequencies and of 100 and higher for radio frequencies. The importance of high values of Q will be pointed out in later discussions.

The impedance of a choke or inductance coil $(R + j2\pi fL)$ is not always linear with frequency and current variations. Thus the resistance component R varies with frequency because of skin effect. If the coil contains an iron core the value of L will vary at low frequencies with the saturation of the core. For the typical magnetiza-

tion curve shown in Fig. 3*a*, the flux in the core will vary directly with H for the range from a to b and L will be constant, but, for the range from b to c, L will be subject to continuous change (nonlinear). With the rise in frequency another form of "skin effect" takes place in the iron core which forces the flux to the outside and makes the iron core less effective. Thus in Fig. 3*b* the flux changes in cylindrical element a of a circular iron core will induce an eddy current in tubes b and c. These eddy currents will oppose the flux change in a and reduce the magnitude of the resulting flux. Similarly, the flux in tube element b will induce eddy currents in tube c which will tend to reduce the flux in both a and b. The result of this action is to produce a flux distribution in the core as shown in the lower part of Fig. 3*b* and to produce a watts loss due to the eddy currents present.

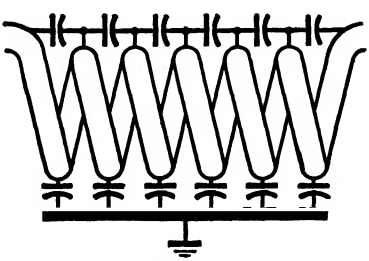


FIG. 4. Distributed capacitance in a coil.

The use of iron cores in choke coils is desirable for raising the Q of the coil but is subject to the limitations arising from hysteresis and eddy currents. Ordinary iron cores (laminated) cannot be used effectively in coils for frequencies above the audio range. Powdered iron cores may be used for frequencies up to 100 megacycles per second because the insulation of the iron particles reduces the eddy currents to very low magnitudes.

Resistors and inductance coils usually possess some distributed capacitance which may introduce troublesome circuit conditions at high frequencies. A single insulated conductor has some capacitance to ground. When wound in the form of a helical coil, a distributed capacitance exists between the individual insulated turns, from turns to ground (Fig. 4), and from layer to layer where two or more layers of insulated conductors exist. Where two or more insulated windings are used as in transformers, there will be a distributed capacitance between windings and between windings and core. Distributed capacitances may produce undesirable resonant effects at high frequencies.

The voltage drop across a capacitance for a fixed frequency is directly proportional to the current flow ($E_c = I_c/2\pi fC$). However, the reactance of a pure capacitance is inversely proportional to the frequency ($X_c = 1/2\pi fC$) and is a nonlinear function with respect to frequency, as shown in Fig. 1. This nonlinear function becomes very useful in resonant circuits. Condensers of practical design have some

inductance and hence do have a resonant frequency. The major source of inductance in a condenser is the self-inductance of its leads plus that introduced by the geometry of its plates (spirals, etc.).

Resonant Circuits. Resonant circuits consist of combinations of inductance L and capacitance C plus some resistance R since inductances and capacitors always contain some ohmic resistance. Typical characteristics of the series-resonant circuit are shown in Fig. 5. Assume that the voltage across the LC combination is held constant and the frequency is varied from zero. The reactances of the inductance X_L and capacitance X_C will vary as previously shown in Fig. 1 and repeated in center view of Fig. 5. The reactance of the series combination Z_T equals $X_L + (-X_C)$ and is depicted by the dotted line. For one value of frequency f_r the inductive reactance X_L equals the capacitive reactance X_C, giving zero reactive impedance for the circuit. This represents the condition of *series resonance* and the frequency f_r is called the *resonant frequency*. The current I flowing in the series circuit for the varying frequency is shown in the lower part of Fig. 5. When the frequency is zero the capacitance blocks the current (direct) because X_C is infinite. As the frequency rises C conducts, and as the frequency of resonance is approached the current rises rapidly, reaching a peak at the resonant frequency (if L and C had zero resistance the current would rise to infinity provided that the impressed voltage were maintained). Beyond resonant frequency tne current falls off rapidly. Since the circuit always has some small resistance the resonant peak rises to some point such as a. If the resistance R of the circuit is increased the resonant peak will be reduced to points such as b and c and the relative width of the peak will increase. Obviously

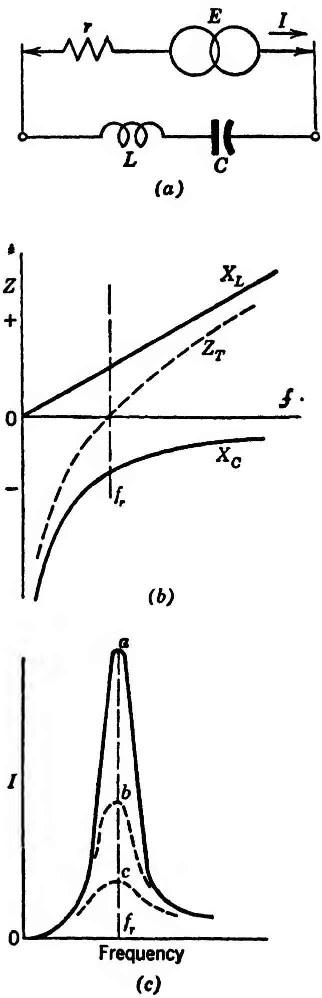


FIG. 5. Series resonance.

the series resonant circuit is highly selective for a narrow band of frequencies and is used where a selective filter is desired.

The resonant frequency depends on the magnitude of L and C and can be made to vary over a wide range by making either L or C variable. Resonance exists when

$$X_L = X_C$$

$$2\pi f_r L = \frac{1}{2\pi f_r C} \tag{2}$$

$$f_r = \frac{1}{2\pi\sqrt{LC}} \tag{3}$$

where f_r is the frequency of resonance.

Since the inductance coil contains some effective or ohmic resistance R, the impedance of the series circuit is

$$Z = R + j\left(L\omega - \frac{1}{C\omega}\right) \tag{4}$$

and at series resonance from equation 2,

$$Z = R \tag{5}$$

This means that at resonance the circuit is purely *resistive* and the line current is

$$I = \frac{E}{R} \tag{6}$$

The voltage drop across L is $L\omega I$ which at resonance from equation 6 must be

$$\text{Voltage across } L \text{ at resonance} = L\omega\frac{E}{R} = EQ \tag{7}$$

Thus with a coil having a high Q it is possible to obtain a high circuit Q and high amplification of voltage across a series resonant circuit. A

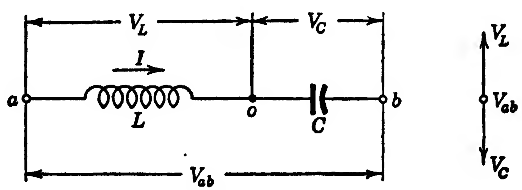


FIG. 6

series LC circuit without resistance and carrying a current I is shown in Fig. 6. At the resonant frequency the voltage drops across the elements will be

$$V_L = IX_L$$

$$V_C = IX_C$$

$$V_{ab} = V_L + V_C$$

and at resonance

$$X_L = -X_C$$

$$\therefore\ V_{ab} = 0$$

Since there is no difference of potential between points a and b, these points may be connected as shown in part a of Fig. 7. The new connection gives a parallel LC circuit having a current circulating within itself. There will be a voltage drop across the parallel circuit equal to IX_L or $-IX_C$, but no current will flow outside the parallel branches. The impedance Z of the parallel circuit by Ohm's law is

$$Z = \frac{V}{I} = \frac{IX_L}{0} = \infty$$

or the impedance Z of the parallel circuit may be calculated thus:

$$Z = \frac{X_L X_C}{X_L + X_C}$$

At resonance,

$$X_L = -X_C$$

hence

$$Z = \frac{-X_L^{2}}{0} = \infty$$

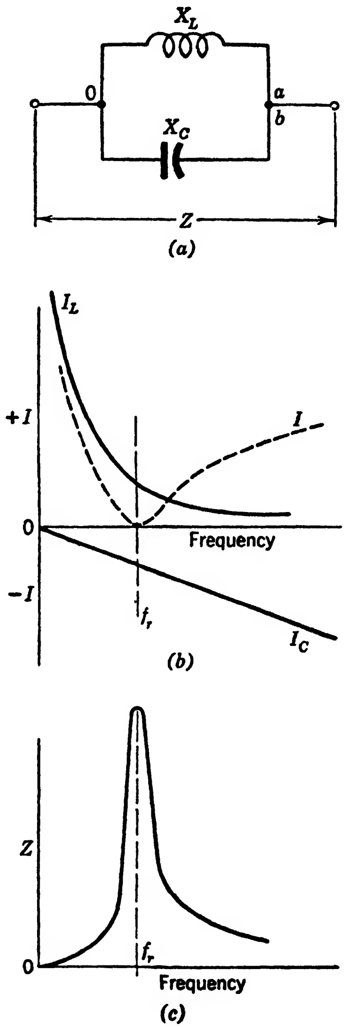


FIG. 7. Antiresonance.

Thus a parallel circuit consisting of pure L and C will offer an infinite impedance at the frequency of resonance. Since the series-resonant circuit offers zero impedance at resonance it is the usual practice to speak of parallel resonance as *antiresonance*. The current in the

parallel branches of L and C and the line current for a variation of frequency are shown in part b of Fig. 7. The variation of the impedance of the parallel LC circuit with frequency has the peaked value shown in part c of Fig. 7.

Some resistance R is present in the inductive arm of the parallel resonant circuit though a good capacitor will show negligible effective resistance. For the parallel circuit of Fig. 8 the parallel impedance may be written

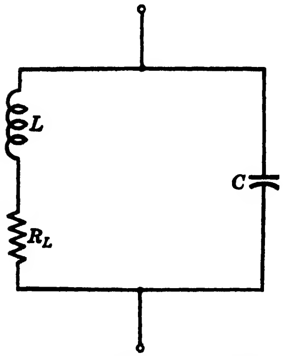


FIG. 8. Series resonance.

$$Z_p = \frac{(R_L + j\omega L)(-j/\omega C)}{R_L + j\omega L - (j/\omega C)} \tag{8}$$

For the condition $j\omega L = j/\omega C$ the denominator reduces to R_L. If Q is large, R_L may be neglected in the numerator, giving

$$Z_p = \frac{(\omega L)^2}{R_L} = QX_L = QX_C \tag{9}$$

and

$$= \frac{Q}{C\omega} = \frac{L\omega}{R_L C\omega} = \frac{L}{R_L C} \tag{10}$$

Equations 9 and 10 show that at anti-resonance *a parallel LC acts as a pure resistance and that the magnitude of the impedance is approximately Q times the reactance of either branch at that frequency.*

The parallel LC circuit is frequently called a *tank* circuit. Its critical or resonant frequency for the anti-resonant condition with large value of Q is calculated by the same formula (equation 3) as for the series resonant circuit. The anti-resonant condition may be produced in an LC parallel circuit for a wide range of frequencies by a variation of the magnitude of L or C as in a series circuit. Whenever a series or parallel LC has its value of L or C adjusted to be in resonance for a given frequency it is said to be *tuned* for that frequency. As such it is called a *tuned circuit*, and tuned circuits play a very important part in electronics.

For industrial applications the Q of a coil is often expressed as volt-amperes divided by the watts, or VI/W. This is a relationship that is derived from the first ratio as follows:

$$Q = \frac{2\pi fL}{R} = \frac{2\pi fLI \times I}{R \times I^2} = \frac{VI}{W}$$

Inductance coils are frequently used in series and parallel resonant circuits where the loss in the capacitor is close to zero and the Q of the coil becomes the VI/W for the resonant circuit.

Resonant circuits have many applications. They may be used for (1) frequency discrimination (i.e., filter action in selecting or rejecting certain frequencies), (2) impedance matching in connecting two units such as generator and load, and (3) voltage amplification.

RC Timing Circuits. The time required for transient current and voltage changes in component circuits consisting of resistors and capacitors is very useful in timing operations. Two combinations of

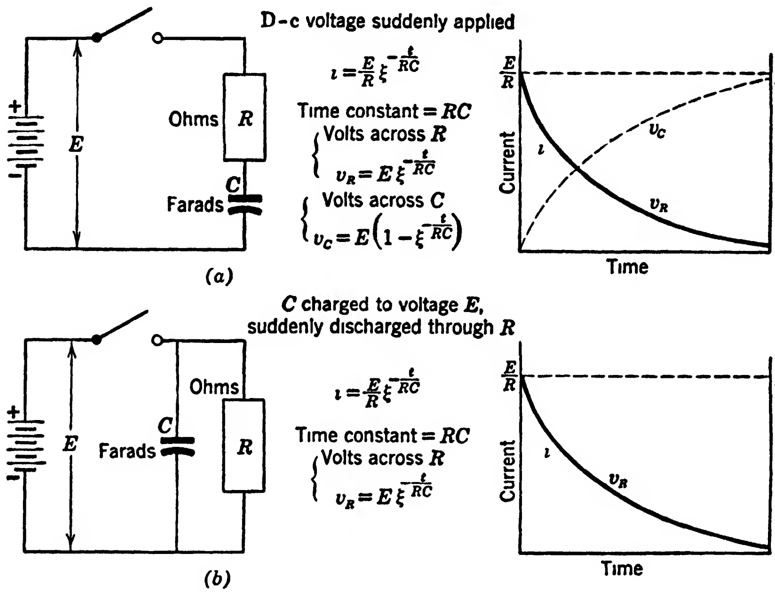


FIG. 9 Assemblies of *RC* timing circuits.

RC circuits together with transient equations and curves are shown in Fig. 9. If a d-c source is used to charge a capacitor in parallel with a resistor and the switch to the supply is opened, the capacitor will discharge at a definite rate determined by the time constant *RC* of the circuit. After an elapsed time equal to the time constant *RC*, the voltage will fall to 37 per cent of its initial value. Similarly, after a second interval equal to *RC*, the voltage will decay by a like ratio; thus this form of timing circuit is very useful. A simple formula for this circuit is time (seconds) equals resistance (megohms) times capacitance (microfarads), for decay to 37 per cent of initial value. Thus, if, in Fig. 9*b*, *R* is 1 megohm, *C* is 1 μf, and *E* is 100 volts, the voltage will fall from 100 to 37 volts in 1 second. Doubling the value of either *R* or *C* will double the time required for the same voltage change. The shape of the voltage decay curve can be controlled by

the substitution of some nonlinear device in the place of a resistor. For example, a pentode properly used in place of R will give a linear decay curve.

Voltage across a capacitor while it is being charged may be used to give another form of timing circuit. In Fig. 10 a resistor and capacitor are connected in series across a source of d-c supply, and a glow tube is placed in parallel with the condenser. When the d-c voltage is applied, current flows through the resistor R into the capacitor and gives it a charge. The voltage across the capacitor [$e_c = q/c$ (charge/capacity)] rises until it reaches the "firing" or discharge value of the glow tube. When the glow tube fires it takes whatever current

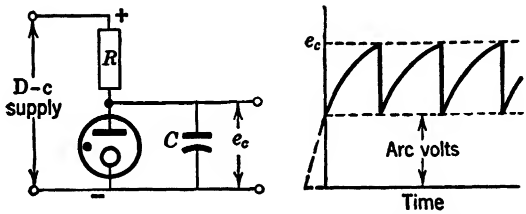


FIG. 10. Simple relaxation oscillator circuit.

is passed by R plus a heavy discharge current from the capacitor. The voltage e_c across the capacitor falls very rapidly until it reaches the extinction point—the minimum ionizing potential across the glow tube.* At this point the glow tube ceases to conduct and the capacitor starts to charge again. When the capacitor voltage reaches the tube firing potential, the process is repeated. This simple circuit, known as a relaxation oscillator, will produce a series of sawtooth voltage timing waves (Fig. 10 right) which are useful in oscilloscopes and similar devices.

Subsequent discussions will cover other applications of RC parallel circuits such as holding bias on the grids of tubes, filters for shunting out r-f voltages and d-c currents, and the feedback for oscillators.

Diode Plate Characteristics. The plate current-voltage characteristic of a vacuum diode was discussed in Chapter III. A portion of this curve is shown in the heavy full-line curve of Fig. 11. This is the static characteristic for voltage readings taken directly across the electrodes of the tube and its curvature follows Childs' equation for the region shown. The dynamic plate resistance r_p for the diode varies with the position on the curve and may be calculated for a

* Operation of a glow tube is explained in Chapter XII.

point as the ratio $\Delta e_b/\Delta i_b$.* For the usual range of operation this resistance r_p does not vary greatly, and, since the value of r_p is usually small compared to the load resistance R_L, it is satisfactory to apply

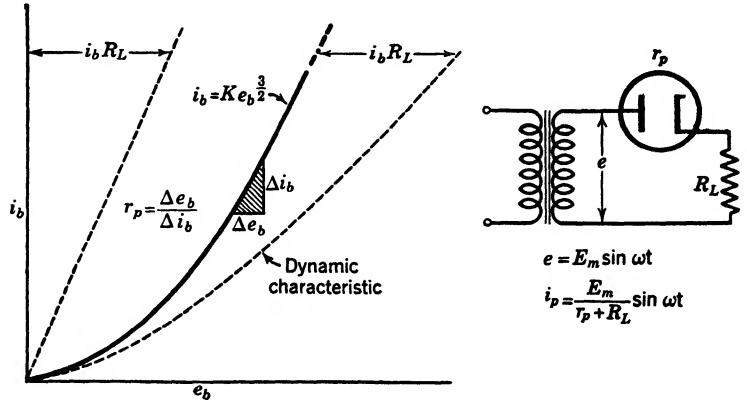


FIG. 11. Static and dynamic characteristics of a vacuum diode.

an average value of r_p in making rectifier calculations. In the complete rectifier circuit the tube resistance r_p appears in series with the load resistance R_L (right side of Fig. 11) and the interest may

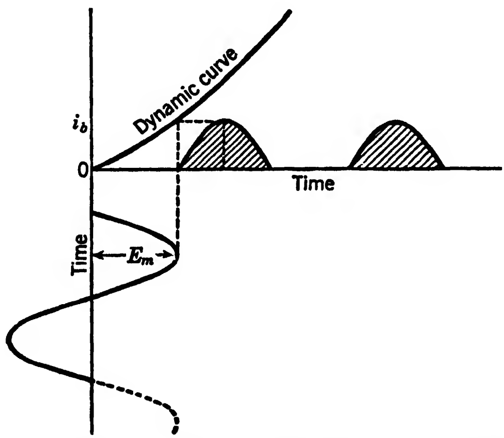


FIG. 12. Graphical solution of diode rectification from dynamic characteristic.

center around the applied a-c voltage e rather than the voltage drop across the tube. In this case the i_bR_L drop across the load may be added to the static curve to give the dynamic characteristic shown by

* Symbols e_b and i_b represent total instantaneous values of plate voltage and plate current.

the dotted line in Fig. 11. The dynamic characteristic will be more nearly linear than the static line. The dynamic curve may be used for providing a graphical representation of half-wave rectification of an a-c wave as shown in Fig. 12. Further analysis and calculations for the process of rectification will be given in a later chapter on rectification.

Triode Plate Characteristic. The dynamic plate-current grid-voltage characteristic for the vacuum triode can be determined from the

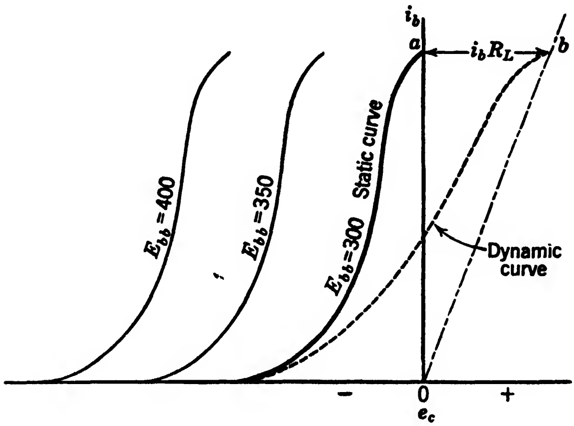


FIG. 13. Dynamic characteristic of a triode.

family of static transfer characteristic curves as illustrated in Fig. 13. That static curve should be selected which corresponds to the desired plate supply voltage (heavy line in figure), and then the grid bias should be selected. To simplify the construction a grid bias equal to the cutoff may be chosen. Then a load triangle $0ab$ may be drawn by adjusting the i_bR_L drop between a and b to the same scale of voltages as the horizontal distance between the static curves on the left. The addition of the i_bR_L drops on the triangle to the static curve gives points for the dotted *dynamic* transfer characteristic. As in the diode, the dynamic curve is more nearly linear than the static curve.

The static transfer characteristic of the triode shown by the dotted line in Fig. 14 is approximately parabolic in form and may be represented by Childs' equation. The dynamic curve may likewise be represented by a parabolic equation. However, the use of the latter equation for predicting the plate output current of a triode when an a-c signal is applied to the grid is not desirable because (1) a mathematical solution is involved; (2) a different equation exists for each value of

load resistance R_L; and (3) more simple graphical methods are available for securing the desired information.

The nonlinear form of the dynamic curve introduces a problem of distortion in the application of the triode as an amplifier. Assume that a load resistor R_L in the simple triode circuit in Fig. 14 results

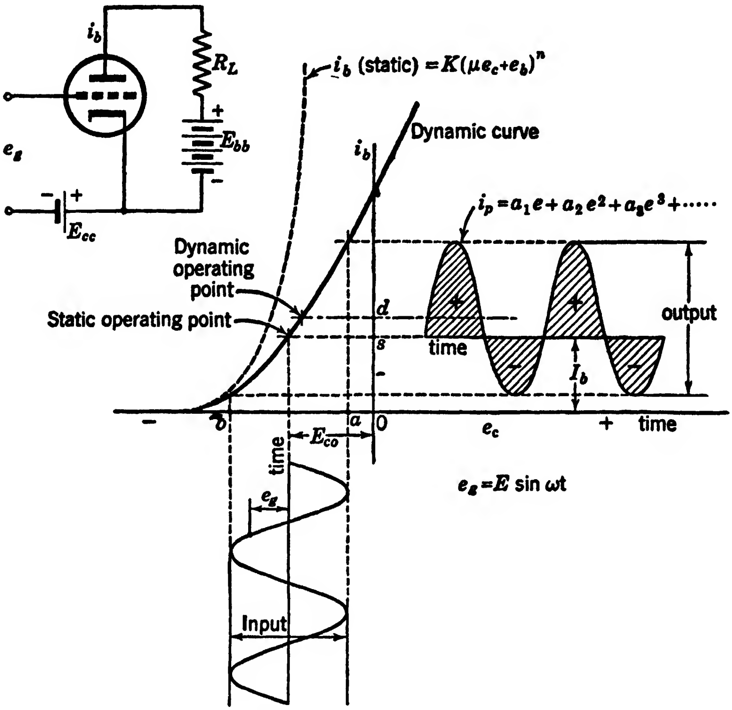


FIG. 14. Graphical analysis of the operation of a triode on its dynamic characteristic.

in the dynamic curve shown by the heavy full line. The negative grid bias of E_{co} acting alone will cause the tube to operate at the static or quiescent operating point and produce a steady plate current with a magnitude 0s. Next assume that an a-c sine-wave signal voltage $e_g = E \sin \omega t$ is applied to the grid of the tube. This signal will cause the resulting voltage on the grid to swing between the limits of points a and b. The construction of dotted vertical and horizontal lines shows that the output plate current will be caused to swing as shown on the right. The positive and the negative loops of the a-c output current as measured from axis through the static operating point are unequal in size. Since the positive loops in the output current of are

larger than the negative, the average of the direct current in the plate current rises from the value $0s$ to some value $0d$. A horizontal axis through d will cut the dynamic curve at a higher point called the *dynamic operating point.* A study of the curve and construction of Fig. 14 shows that this inequality in the positive and the negative current loops is due to the nonlinearity of the dynamic curve since with a linear curve they would be identical. This inequality of the alternating current loops is called *amplitude distortion* or nonlinear distortion, and such distortion is very undesirable where the tube becomes a link in a circuit chain for reproducing music and other sounds.

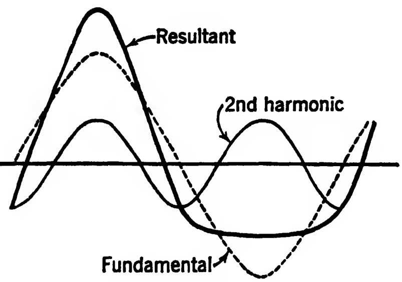


FIG. 15. Nonsymmetrical resultant wave from combination of the fundamental and second harmonic.

Amplitude distortion is marked by the presence of harmonics (multiples of fundamental) in the output current. This statement is verified pictorially in Fig. 15 in which a fundamental is combined with its second harmonic to produce an unsymmetrical wave which is characteristic of amplitude distortion.

The presence and the measurement of harmonics in the output current of a vacuum tube may be determined by analytical and graphical methods. The alternating component of the output current i_p shown in Fig. 14 may be represented by the following general form of power series

$$i_p = a_1 e + a_2 e^2 + a_3 e^3 + \cdots \tag{11}$$

where e is the signal being applied to the grid. If a sinusoidal voltage $E_g \sin \omega t$ is assumed to be applied to the grid, equation 11 becomes:

$$i_p = a_1 E_g \sin \omega t + a_2 E_g{}^2 \sin^2 \omega t + a_3 E_g{}^3 \sin^3 \omega t + a_4 E_g{}^4 \sin^4 \omega t + \cdots \tag{12}$$

The substitution of equivalent trigonometric functions for the $\sin^2 \omega t$, $\sin^3 \omega t$, and $\sin^4 \omega t$, followed by a collection of like terms, gives an equation which reduces to the form

$$i_p = K + H_1 \sin \omega t - H_2 \cos 2\omega t - H_3 \sin 3\omega t + H_4 \cos 4\omega t + H_5 \sin 5\omega t \cdots \tag{13}$$

This equation shows that any form of current wave for i_p can be produced fundamental wave plus a series of harmonics of suitable

amplitude (H factor) and a constant or d-c component of magnitude K. The factor K is the d-c component represented by the distance sd on Fig. 14. The total current flowing in the plate circuit i_b is that represented by equation 13 plus the quiescent direct current I_b

$$i_b = I_b + K + H_1 \sin \omega t - H_2 2 \cos \omega t - H_3 \sin 3\omega t + H_4 \cos 4\omega t + H_5 \sin 5\omega t \cdots \quad (14)$$

In the application of vacuum tubes it is necessary to know the amount of amplitude distortion, and that means the magnitude of the

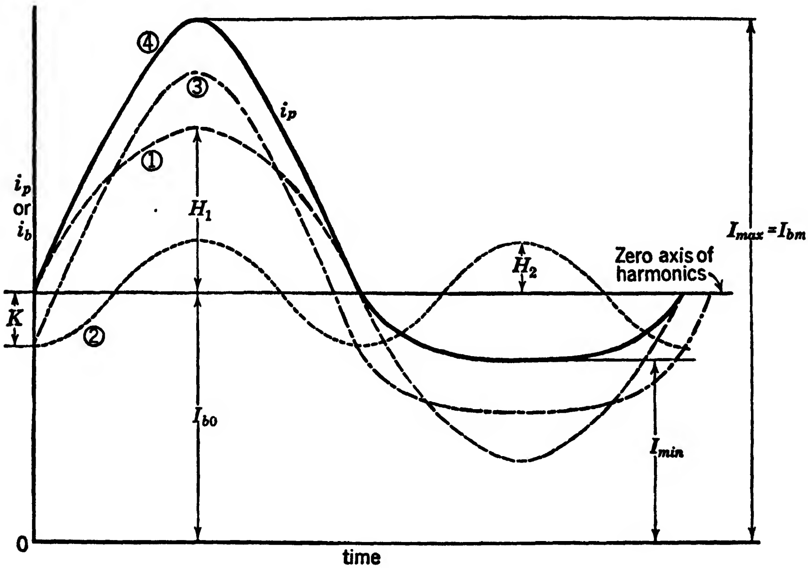


FIG. 16. Graphical representation of the components in a nonsymmetrical plate current.

H factor of the harmonics. The H factor may be determined by analytical or graphical methods or by a combination of both. A simple method of determining the H factors for the fundamental and second harmonic where the other harmonics may be considered negligible is illustrated in Fig. 16. The heavy line of this figure is assumed to be a lopsided or unsymmetrical current-output curve which has been formed by combining a fundamental sine wave of amplitude H_1 with a second harmonic of amplitude H_2 and adding the d-c plate component to give the total plate current i_b. To understand these curves note the following points. The steady or quiescent plate current (no grid signal) is I_{bo}. When the fundamental sine wave of curve 1 (ampli-

tude H_1) is combined with its second harmonic of curve 2 (amplitude H_2), the resultant wave form is represented by curve 3. The combination of the fundamental and second harmonic produces a d-c shift component of some magnitude K. Hence it is necessary to elevate curve 3 by the magnitude K, bringing it into position shown by curve 4. Now, referring to the ordinate of zero time, it will be observed that the upward shift K is equal to the amplitude of the second harmonic H_2.

$$K = H_2 \tag{15}$$

At the ordinate corresponding to one-fourth cycle of the fundamental frequency wave the plate current reaches a maximum $I_{\max}$ where

$$I_{\max} = I_{bo} + H_1 + H_2 + K \tag{16}$$

and at the three-fourths cycle of the fundamental frequency wave the plate current reaches a minimum where

$$I_{\min} = I_{bo} - H_1 + H_2 + K \tag{17}$$

The solution of equations 15, 16, and 17 reduces to

$$H_1 = \tfrac{1}{2}(I_{\max} - I_{\min}) \tag{18}$$

$$H_2 = \tfrac{1}{4}(I_{\max} + I_{\min}) - \tfrac{1}{2}I_{bo} \tag{19}$$

$$K = \tfrac{1}{4}(I_{\max} + I_{\min}) - \tfrac{1}{2}I_{bo} \tag{20}$$

These equations represent the constants for the first three terms of equation 13. A similar line of reasoning can be applied to determine the formulas for the constants for the third and higher harmonics. In most design problems all harmonics above the third are small and may be neglected. In triodes the second harmonic is the most important cause of distortion and often the effect of other higher harmonics may be neglected because of their small magnitude. The magnitude of the terms $I_{\max}$, $I_{\min}$, and I_{bo} for calculations involving the preceding equations may be taken from developments from the dynamic curve, as shown in Fig. 14, but the usual and convenient practice is to secure them from the dynamic load line as will be explained in a succeeding chapter.

If a small a-c signal voltage is introduced into the grid circuit, the tube will operate over a small section of the dynamic curve which is

more nearly linear and the distortion will be correspondingly reduced. With the increase of the signal strength and the swing of the grid voltage, the amplitude distortion and the magnitude of the harmonics increase. (It is obvious that the operation of the triode over any section of the dynamic curve which is nearly linear will give low distortion.

Problems

1. A choke coil takes a current of 250 ma and consumes 3.0 watts at 5000 cycles. What is the effective resistance of the coil? If the impressed voltage across the coil is 120 volts rms, what is the Q and L of the coil for the given set of conditions?

2. A coil with an air core has an effective resistance of 28 ohms and inductance of 48 mh at 2500 cycles. What is the Q of the coil?

3. A series RLC circuit has a resistance of 65 ohms, a capacitance of 1.0 μf and an inductance of 15.0 mh. What is the resonant frequency and the impedance of the circuit at resonance?

4. A parallel LC circuit consists of a choke coil having an L of 2.5 mh with a resistance of 2.5 ohms, and a capacitance of 4.0 μf. Calculate (1) the frequency at anti-resonance (2) the Q of the circuit, and (3) the impedance of the circuit.

5. A tuned parallel circuit has a Q of 70 at its resonant frequency of 80 KC. If the L of the choke is 2.0 mh, what is the impedance of the tuned circuit at resonance?

6. A parallel tuned circuit has a Q of 85 and a capacitance of 0.1 μf at the resonant frequency of 150 KC. What is the impedance of the tuned circuit at resonance?

7. In the RC timing circuit of Fig. 9*b* having a resistance 1.0 megohm and an impressed voltage of 100, the voltage across the resistor drops to 37 volts in 3.0 seconds. What is the magnitude of C? What is the magnitude of i at this point?

8. Design an RC circuit like Fig. 9*b* where the voltage across R will fall to 13.7 per cent of its initial value in 5 seconds.

9. In Fig. 9*a*, R has a value of 0.5 megohm and C is 4.0 μf with E equal to 115 volts. How long will it take for the voltage across the condenser to rise to 72.5 volts after S is closed? What will be the magnitude of the current 4 seconds after the switch is closed?

10. The quiescent current in the plate circuit of a triode amplifier is 30 ma. For a constant-frequency sine-wave input to the grid the plate current has a maximum value of 60 ma and a minimum of 20 ma. Calculate the H_1, H_2, and K factors.

Chapter VI

VACUUM-TUBE AMPLIFIERS

Amplification. The vacuum triode, tetrode, pentode, and beam-power tube are amplifying tubes. Their association with suitable circuits constitutes one type of amplifier, and as such they have revolutionized the art of electrical communication, measurement, and control. The function of the amplifier is to increase or magnify a very weak signal until it is capable of controlling sufficient electric power for producing light, sound, or mechanical work. An amplifier does not create energy but controls sources of electric energy to give outputs that follow the form of the original input signal. An amplifier performs its function in one stage or in a series of stages or steps.

The individual stage of an amplifier may be designed for (1) voltage amplification, (2) current amplification, or (3) power amplification. Voltage amplification signifies that the output voltage across the load is greater than the voltage in the input signal and that the power output is small. Current amplification is used when the load on the amplifier requires a relatively large current rather than voltage for operation. Power amplification means that the output power (volts times amperes times power factor) is large compared to that of the input signal. Frequently a complete amplifier consists of one or more stages of voltage amplification succeeded by a final stage of power amplification. Any amplifying tube may be provided with a circuit to give either voltage, current, or power amplification. However, the maximum allowable output of a tube is determined by its size, construction, and the heat dissipation of its plate. Hence design practice usually employs a small tube of low plate dissipation for voltage amplification and a larger tube of high plate rating for the power-amplification stage. The over-all power gain in an amplifier of several stages may be very large—approaching infinity—because the input to the first stage of voltage amplification may approach zero, whereas the energy output of the final stage may be large. The student should note that the phenomenon of amplification is essentially the same whether the desired output is power, voltage, or current.

Amplifier Distortion. The function of an amplifier is to increase the energy level of the signal applied to the input terminals without changing its wave form or frequency spectrum. If the output of the amplifier is not an exact replica of the input, *except in magnitude,* distortion

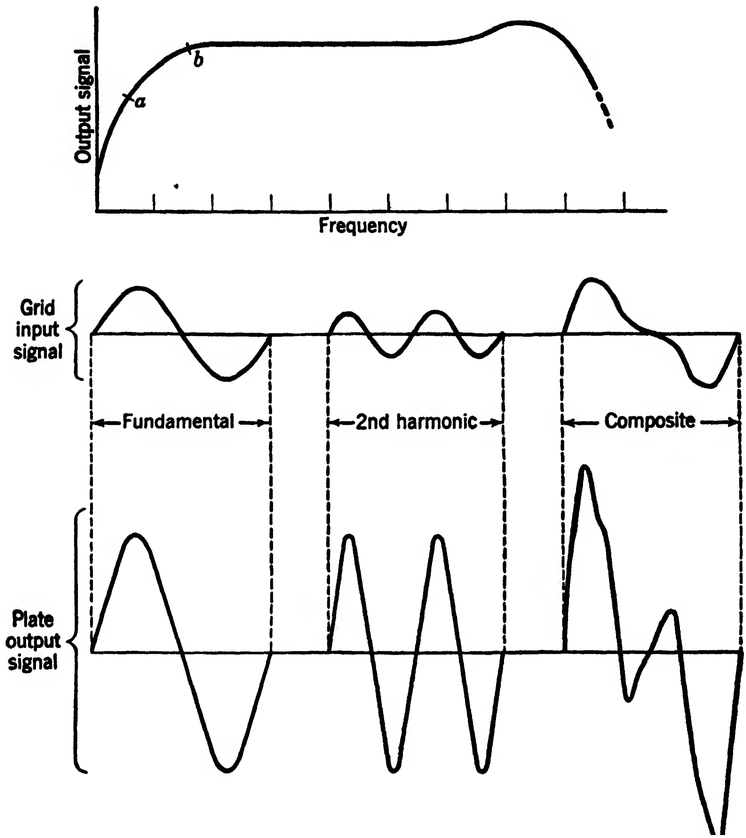


FIG. 1. Graphical representation of frequency distortion in an amplifier.

is present. The types of distortion are called amplitude, frequency, and phase.

Amplitude distortion exists when the magnitude of the output of the amplifier is not directly proportional to the input. This type of distortion occurs when the dynamic transfer characteristic of the amplifier is not a straight line as explained in the preceding chapter. Amplitude distortion is maximum (assuming proper tube bias) when an amplifier is delivering outputs near or in excess of its maximum rating.

Frequency distortion is caused by an unequal amplification of the different frequency components in a signal. This is illustrated by the gain versus frequency characteristic of an amplifier shown in Fig. 1. Frequency distortion for the fundamental at point *a* and its second harmonic at point *b* is illustrated in the middle and the lower views of Fig. 1. Frequency distortion is usually the result of the design of the associated circuit rather than an inherent property of the amplifying tube.

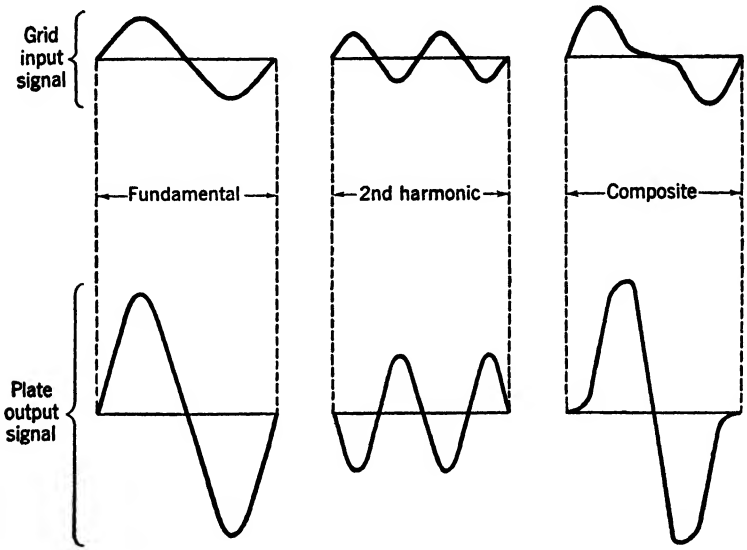


FIG. 2. Phase distortion in an amplifier.

Phase distortion results when the phase relation between the sinusoidal components of a complex signal is changed by the amplifier. Hence the time of transmission through the amplifier circuit differs for various frequencies. Phase distortion is illustrated in Fig. 2.

The amount and kind of distortion that can be tolerated varies with different types of amplifiers. Amplitude distortion must be reduced to the minimum in audio-frequency amplifiers as the human auditory system is very sensitive to the discordant tone combinations that result. The presence of frequency distortion in audio amplifiers is not so objectionable, provided that the frequency band width of the amplifier is adequate to pass an intelligible signal and the gain over the pass band is reasonably uniform. Ordinary speech is transmitted without loss in intelligibility if the pass band is limited to a range of 200 to 2500 cycles. The average individual fails to detect the loss in

quality of transmitted music when the pass band is limited to a range of 100 to 5000 cycles. Human beings are incapable of detecting phase distortion by ear unless the time-delay variations over the pass band are many times greater than those found in audio amplifiers.

The distortion problem in the video amplifiers used in television and radar systems is the reverse of that for audio amplifiers. Video amplifiers must amplify the input wave form with minimum change in shape. Any complex periodic wave can be resolved into sinusoidal components with finite amplitudes and phase displacements.* The most important factor in determining the wave shape is the phase displacement between the components. Therefore, phase distortion must be held to the minimum in video amplifiers. Frequency distortion usually accompanies phase distortion, but its distorting effect on the wave shape is much less. The importance of minimizing amplitude distortion depends on the type of wave form being amplified. Considerable amplitude distortion can be tolerated in a television video amplifier as its major effect is to reduce the contrast in either the highlighted or lowlighted areas, or both, of the picture. Often amplitude distortion is introduced into a television system to reduce the contrast range.

Phase distortion or time delay is an important problem in amplifiers used in control systems such as speed regulators because the time delay may cause instability or "hunting" in the controlled mechanism.

Factors Limiting Amplifier Performance. Voltage amplification of more than one million is possible in a well-designed amplifier circuit. The limit of maximum usable amplification is reached when spurious currents generated within the amplifier approach in magnitude the desired signal current. These unwanted currents are called "noise currents" and can be divided into:

(*a*) "Shot noise"—generated by the random and varying emission of electrons from the cathode.

(*b*) "Thermal noise"—due primarily to the movement of free electrons in the grid circuit (usually in grid leak resistor).

(*c*) "Partition noise"—present in tetrodes and pentodes due to slight changes in the number of electrons falling on the screen grid as a result of random variations in transit to the plate.

(*d*) Microphonics—caused by vibration of the elements within the tube or of other components and wiring in the circuit.

* Reddick and Miller, *Advanced Mathematics for Engineers,* Chap. V, Fourier Series, John Wiley & Sons, 1947.

(*e*) Hum and other noises—caused by stray electrostatic or electromagnetic fields and by improperly filtered supply voltage.

A study of the process by which electrons are released from metals reveals that this emission is completely at random. Hence at any one instant, more (or fewer) electrons may break away from the attraction of the cathode than at some other instant. If a large space charge exists in the tube, fluctuations in the current reaching the plate will be greatly reduced because the space charge acts as a reservoir of electrons. In the absence of space charge, plate current must vary as the emission varies. Shot noise can best be minimized, therefore, by using a cathode with ample emission and operating the control grid as far negative as possible.

In any conductor loosely bound electrons in the outer shells of the atoms are continually changing places with those of other atoms. This motion is at random so that at any instant more electrons may be moving in one direction than in the other.

Since this noise voltage varies with temperature, it is called thermal noise and is known as the Johnson effect. Because the energy of thermal and shot noise is uniformly distributed over the entire frequency range, one of the best ways in which to reduce their effect is to limit the pass band of the amplifier to include only those frequencies actually needed. The grid circuit resistance also can be reduced, but when this causes a corresponding reduction in desired signal no improvement in signal-to-noise ratio will result.

Microphonics can be controlled best by tubes of rugged construction where the elements are small and rigidly supported. A heavy shield around the tube usually keeps sound pressure waves from vibrating the elements. A cushion mounting for the tube usually helps, and sometimes an entire multiple-stage amplifier must be shock-mounted. In a high impedance load circuit followed by a large amount of amplification, the vibration of components such as condensers, resistors, and wiring may produce spurious voltages many times greater than the desired signal. Leads should be short and rigid with all components anchored securely in place. In extreme instances it may be necessary to "pot" the entire assembly in wax.

Amplifiers operating from a rectified a-c supply may pick up stray hum if suitable precautions are not observed. All plate voltage supply ($B+$) leads should be well filtered. When alternating current on the filaments of low-level stages are used, the cathodes should be connected directly to "ground" (the common side of the circuit) or the voltage between heater and cathode should be adjusted to minimize

hum due to heater-cathode leakage and capacity. If a cathode resistor must be used for grid bias, a large by-pass condenser should be connected from cathode to ground. Audio transformers should have balanced windings and thick well-fitted iron shields. Power transformers and filter chokes should be kept well separated from the audio transformers. A heavy copper band placed around the power transformer will act as a shorted turn and materially reduce stray leakage flux. Electrostatic fields and circulating currents in the amplifier chassis often cause hum, which can be eliminated by proper placement of parts, shielding, and correct location of ground points.

Factors which limit the highest frequency that can be amplified by a vacuum tube are lead inductance, interelectrode capacity, and electron transit time. These may be minimized by proper design of the tube and circuit.

Gain as a Logarithmic Function. The human ear responds to sounds of different intensities in a logarithmic manner. For this reason it is desirable to show the gain (or loss) of a circuit as a logarithmic function. The standard unit which expresses this relationship is the bel, so named in honor of Alexander Graham Bell. A convenient practical unit is the "decibel" (one-tenth of a bel), which is about the smallest difference in power level the ear can detect. The decibel is usually abbreviated to db and is defined by the equation

$$\text{db} = 10 \log_{10} \frac{P_2}{P_1} \tag{1}$$

where P_1 is the input power and P_2 is the output power of the circuit under consideration.

If the output power is less than the input, the circuit has a net loss in gain which is designated by a negative sign. A gain of unity will be equivalent to 0 db, since the logarithm of 1 is 0. It is preferable to specify changes in output power as "db gain" or "db loss" rather than by positive or negative signs.

When the resistances in which powers P_1 and P_2 are developed are equal, then

$$\text{Decibels (db)} = 10 \log_{10} \frac{E_2^2/R}{E_1^2/R} = 20 \log_{10} \frac{E_2}{E_1} \tag{2}$$

and

$$\text{Decibels (db)} = 10 \log_{10} \frac{I_2^2R}{I_1^2R} = 20 \log_{10} \frac{I_2}{I_1} \tag{3}$$

A conveniently arranged table or a curve like that in Fig. 3 can be used to advantage when power, current, and voltage ratios are converted to db.

It has become rather common to speak of the over-all voltage gain of an amplifier in terms of decibels without regard to any difference between input and output resistance. Since gain in decibels is defined as a function of the *power ratio,* this usage is not strictly cor-

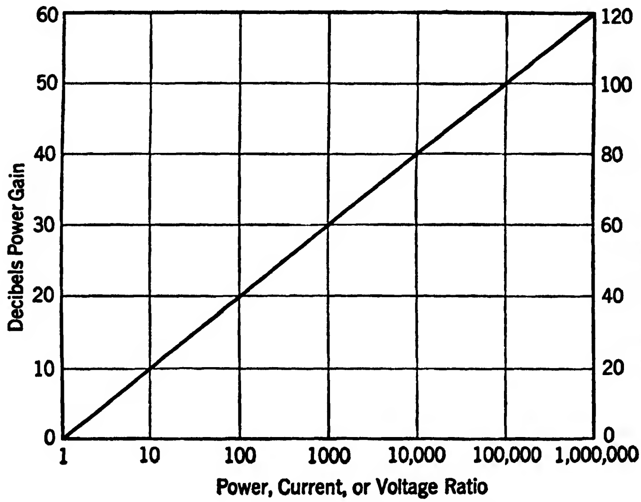


Fig. 3. Logarithmic gain in amplifiers.

rect, and it should always be clearly stated when a *voltage ratio* only is considered.

Often the output of a circuit is compared to some arbitrary reference level. Six milliwatts was chosen by the telephone industry since that was the normal output of a standard telephone "repeater" tube. Based on this reference level, a circuit delivering 3 mw is said to have an output of -3 db. An output of 6 watts is $+30$ db. Other reference levels commonly used are 10 mw and 12 mw. In an attempt to standardize on one power reference level, the term "volume unit" (vu) has been adopted in place of decibel. This unit has the same defining equation as the decibel but is always based on a reference level of 1 mw, a reference resistance value of 600 ohms, and with definitely specified characteristics for the measuring instruments.*

* *Proc. I.R.E.,* January 1940, "A New Standard Volume Indicator and Reference Level," by H. A. Chinn, Dr. Garnett, and R. W. Morris.

Decibels are especially convenient when the total gain or loss in a system is being calculated since it is only necessary to add algebraically the figures for each part of the circuit. For example, in a certain telephone circuit the line may have a loss of 12 db, the equalizer a loss of 6 db, and the repeater a gain of 20 db. Total gain is $-12 - 6 + 20 = +2$ db.

The Vacuum Tube as a Negative Grid Amplifier. Figure 4 is the circuit of a triode amplifier with a resistance load. Tetrode and pentode tubes are connected like triodes except that the screen grid must be supplied with the proper positive voltage, usually less than E_{bb},

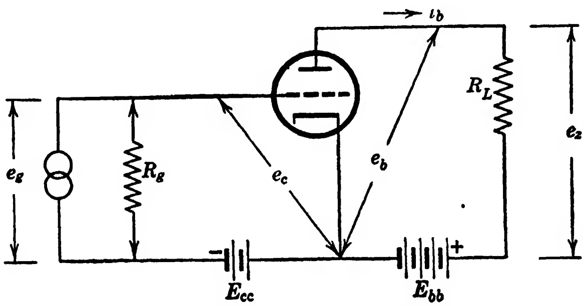


FIG. 4. Basic amplifier circuit.

and the suppressor connected to the cathode. The screen-to-cathode impedance at signal frequency must be low to prevent degeneration and loss in gain. For simplicity, batteries are shown as the plate and grid supplies, whereas the filament circuit has been omitted. The principles of operation to be discussed will, of course, be the same with heater-type tubes operated from a rectified a-c supply. Standard nomenclature covering amplifier circuit parameters has been adopted by the Institute of Radio Engineers (Standards on Electronics, 1938) and is summarized in the table opposite page 1.

When a signal voltage e_g is applied to the grid, the plate current i_b follows the variations in grid voltage. If the grid potential becomes more negative, i_b decreases; if the grid swings in the positive direction, plate current increases. Since the plate current flows also through the plate load resistor R_L, any variation in current will cause a change in the i_bR_L drop. Proper choice of μ, r_p, and R_L will cause the plate-voltage variation to be many times that of the original signal, and voltage amplification is produced.

In normal operation where the grid is held negative with respect to the cathode, the current flowing to a grid i_c is negligible. Therefore,

if the resistance of R_g is high, little power will be drawn from the signal source and the *power gain* in the stage will be very great. The source for this output power is the plate supply E_{bb} with the amplifier acting as a converter changing d-c power into a-c signal power. Referring to a hydraulic analogy, the tube acts like the valve in a water pipe where a small force exerted on the handle of the valve may control the flow of an enormous volume of water. For this reason vacuum tubes are sometimes called "valves."

The phase relationship of the various voltages and currents in an amplifier stage produced by a sine-wave signal is shown in Fig. 5.* The a-c grid signal e_g is shown at the zero axis. This signal is shifted by the grid bias to an axis about E_{co} giving a resultant instantaneous value such as e_c. With zero a-c signal on the grid, the plate current assumes its quiescent value indicated by I_{bo}. With the application of an a-c grid signal the plate current follows the grid, giving a total instantaneous current of value i_b, a varying a-c component of i_p, and an average value of I_{ba}.† If the swings of a-c plate current take place over a linear section of the dynamic curve, the quiescent value I_{bo} and the average value I_{ba} will be the same, but for the nonlinear condition the magnitude of the average value I_{ba} will be greater for reasons discussed in the preceding chapter. For a zero grid signal there is a constant quiescent voltage E_{bo} across the tube. With the introduction of the a-c grid signal the plate voltage varies, having instantaneous total values such as e_b and instantaneous varying component values of e_p. The axis of the a-c plate voltage shifts for operation on nonlinear dynamic curves. The important thing to be noted in this analysis of phase relations is that the plate current is in phase with the grid signal *but that the plate voltage variation is 180 degrees out of phase with the grid signal.* The instantaneous output voltage e_z across the load is the drop produced by the current flowing through the load resistance R_L, or Z_L where reactance is present.

The operation of the negative grid amplifier can be analyzed by two methods. In one method the assumption is made that the tube is a linear device, i.e., operation occurs with a linear dynamic characteristic. This method employs the concept of a simple equivalent circuit and serves well for signals of small amplitude or where operation occurs over a linear section of the dynamic curve. The second method of analysis accepts the nonlinearity of the tube characteristic and em-

* See table opposite page 1 for review of letter symbols.

† Often given as I_{bs}.

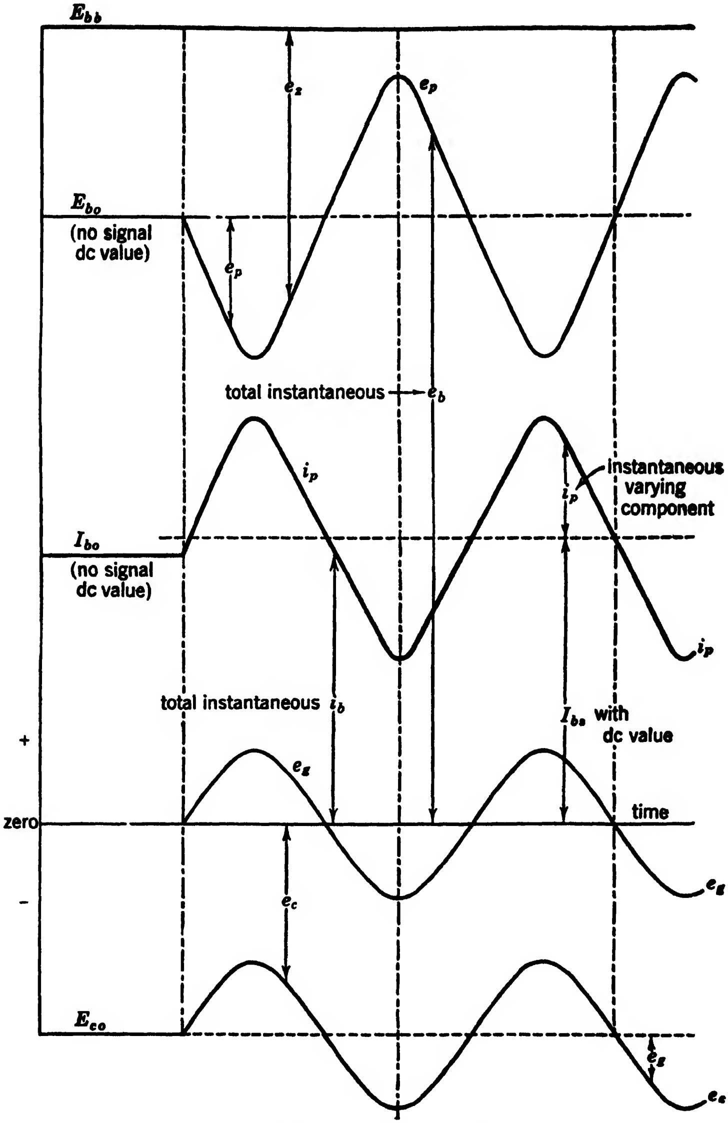


FIG. 5. Current and voltage phase relations in a triode amplifier.

ploys graphical solutions with load lines placed on a family of plate characteristics of the tube.

Equivalent Circuits. The calculation of the gain for a single stage of a vacuum-tube amplifier can be greatly simplified by the use of the equivalent circuit shown in Fig. 6. In part *a* of this figure the plate circuit of the vacuum tube is replaced by a generator having an emf of μe_g and a resistor having the same value as the internal or dynamic

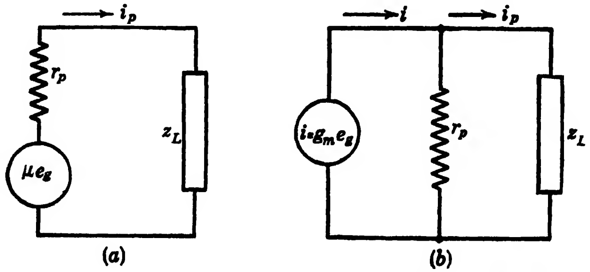


FIG. 6. Equivalent circuits for amplifiers (*a*) constant-voltage form, (*b*) constant-current form.

plate resistance r_p of the tube. This generator circuit is connected to the plate load Z_L. If the load is a pure resistance R_L, the plate current will be:

$$i_p = \frac{\mu e_g}{r_p + R_L} \tag{4}$$

where e_g and i_p are the instantaneous values of the varying components. The current considered is the *electron current* flowing from cathode to plate and thence through the load impedance. Such a current will produce a *voltage rise* in the direction of current flow. Since the voltages in the plate circuit are 180 degrees out of phase with the signal voltage e_g (assumed positive), the varying component of load voltage across the load resistor will be:

$$e_z = -i_p R_L = -\frac{\mu e_g R_L}{r_p + R_L} \tag{5}$$

Dividing both sides by e_g gives the voltage gain G:

$$G = \frac{e_z}{e_g} = -\frac{\mu R_L}{r_p + R_L} \tag{6}$$

If the load contains inductive or capacitive components as well as resistance, the above methods may be used by handling the quantities in complex notation, giving the following equations.

$$Z_L = R_L + jX_L$$

$$G = -\frac{\mu Z_L}{r_p + Z_L} = -\frac{\mu(R_L + jX_L)}{r_p + R_L + jX_L} \tag{7}$$

The equivalent circuit of Fig. 6*a* is the "constant voltage generator" form since a constant signal voltage e_g was assumed. It is sometimes convenient to start with a "constant current generator" form as shown in Fig. 6*b*. Since $\mu = g_m r_p$, where g_m is the mutual conductance of the tube, equation 4 can be written

$$i_p = \frac{g_m r_p e_g}{r_p + R_L} \tag{8}$$

Also

$$e_z = -i_p R_L = -\frac{g_m r_p e_g R_L}{r_p + R_L} \tag{9}$$

The load voltage is therefore that which results from a current $-g_m e_g$ flowing through a resistor equal to r_p and R_L in parallel, and the gain is

$$G = \frac{e_z}{e_g} = -\frac{g_m r_p R_L}{r_p + R_L} \tag{10}$$

Here again the negative sign indicates 180 degrees phase reversal.

If r_p is very high compared to R_L (as is usually true with tetrodes and pentodes), equation 10 reduces to the very simple and convenient form

$$G = -g_m R_L \tag{11}$$

The constant-current generator type of equivalent circuit shown in Fig. 6*b*, which utilizes equations 10 and 11 for calculating voltage gain, is useful where the plate load consists of several components in parallel as is generally true in multistage amplifier circuits.

The solution found through the use of equivalent circuits will be exact provided that the assumed values of μ, g_m, and r_p remain constant over the entire operating range. For small signals this is very nearly true.

The application of equivalent circuit formulas to amplifiers having a variety of couplings between stages requires a careful understanding of the nomenclature for the load circuit components. The system and nomenclature to be followed is indicated in Fig. 7. The resultant impedance into which a given amplifier stage feeds will be considered as R_L when it is resistive and Z_L when it is complex. The resistance load on the plate circuit ahead of the coupling condenser will be termed R_v.

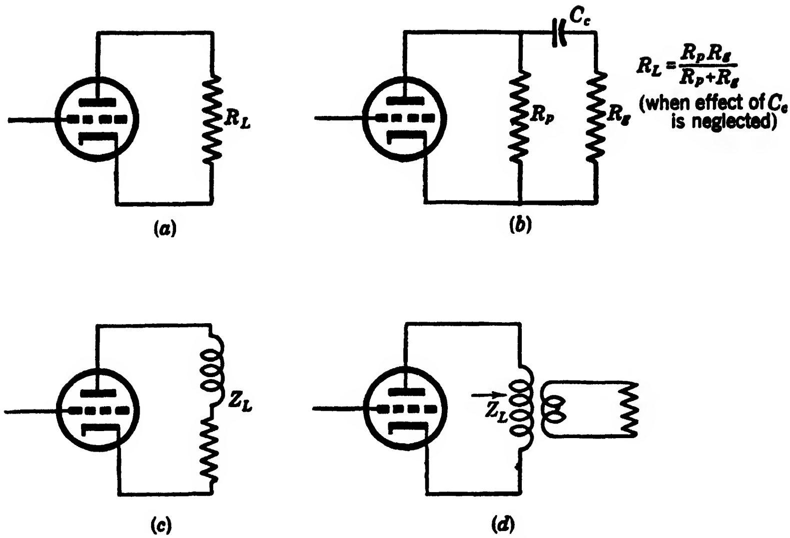


FIG. 7. Letter symbols for various loads in equivalent circuits.

Load Lines. The operation of a vacuum-tube amplifier may be determined graphically also. In this method a "load line" for a given load is drawn on the family of plate-voltage plate-current curves. These curves can be obtained from the tube manufacturer, or they may be determined by experiment in the laboratory. The load line is the locus of the i_b-e_b points for any given value of load R_L and has a slope equal to $-1/R_L$. In a triode operated with a pure resistance load (Fig. 4), the load line is determined as follows (see Fig. 8):

(*a*) Locate point A at the power supply voltage E_{bb} and at $i_b = 0$.

(*b*) Determine point B by taking E_{bb}/R_L. This is the current that would flow if there were no drop through the tube and the entire supply voltage appeared across the load resistor.

(*c*) Connect points A and B with a straight line.*

The quiescent point P is the intersection of the load line and the particular grid-bias curve being considered. Thus, in Fig. 8 with no signal applied and a grid bias of -4 volts, the plate current will be 3.6 ma and the voltage 260. The total power drawn from the $B+$ supply will be $0.0036 \times 400 = 1.44$ watts, of which $0.0036 \times 260 = 0.94$ watt will be dissipated by the plate of the tube and the remainder by the load resistance.

* If point B falls off the curve sheet, assume some fractional part of E_{bb} and divide by R_L. Then plot a point corresponding to the assumed value of voltage and the calculated current value. A line drawn through A and this point will be the load line.

Upon the application of signal to the grid, the instantaneous ***operating point*** will move along the load line in accordance with the instantaneous grid voltage. Assuming that the grid signal swings from the quiescent point of -4 volts to -2 volts and then to -6 volts, the maximum swing or peak-to-peak value is 4 volts. Vertical and horizontal projections from the intercepts of these grid voltages on the load line will show that, for this peak-to-peak swing on the grid, the plate voltage e_b has changed from 210 to 305 volts and the plate

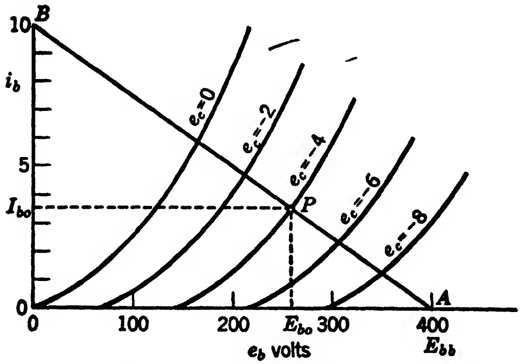


FIG. 8. Determination of amplifier performance with the load line.

current i_b has varied from 4.8 to 3.3 ma. This concept may be stated in equation form and applied where nothing higher than the second harmonic is present, as follows.

	PLATE VOLTAGE	PLATE CURRENT
Peak-to-peak value	$E_{b\,\max} - E_{b\,\min}$	$I_{b\,\max} - I_{b\,\min}$
Peak value	$\dfrac{E_{b\,\max} - E_{b\,\min}}{2}$	$\dfrac{I_{b\,\max} - I_{b\,\min}}{2}$
Rms value	$\dfrac{E_{b\,\max} - E_{b\,\min}}{2\sqrt{2}} = E_p$	$\dfrac{I_{b\,\max} - I_{b\,\min}}{2\sqrt{2}} = I_p$

Thus

$$\text{Peak power} = \frac{(E_{b\,\max} - E_{b\,\min})(I_{b\,\max} - I_{b\,\min})}{4} \tag{12}$$

$$\text{Rms power} = \frac{(E_{b\,\max} - E_{b\,\min})(I_{b\,\max} - I_{b\,\min})}{8} \tag{13}$$

or

$$\text{Rms power} = I_p^2 R_L = I_p E_p$$

$$\text{Gain} = G = \frac{E_{b\,\max} - E_{b\,\min}}{E_{g\,\max} - E_{g\,\min}} \tag{14}$$

If distortion is not present the average plate current remains constant with and without signal. The total power drawn from the plate supply averaged over a complete cycle is a constant, and any a-c power which is delivered to the load must be subtracted from the steady-state plate

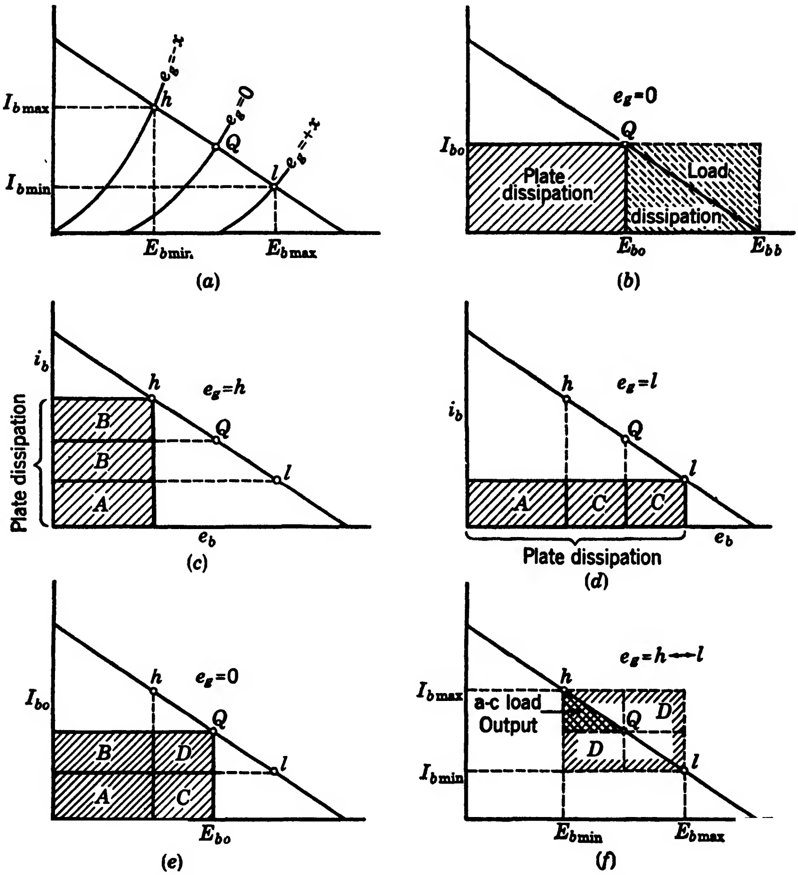


FIG. 9. Graphical representation of the losses in a tube under operating conditions.

dissipation. In other words, for class A operation, less power is wasted in heating the plate of the tube when a signal is applied to the grid. This statement can be demonstrated semigraphically, though not rigidly, by the six steps shown in Fig. 9. A tube having the plate characteristics and load line given in Fig. 9*a* is assumed to be operating about point Q which is midway between the high h and low l swings of its grid potential (zero harmonic distortion). When oper-

ating in its quiescent state with zero signal applied to the grid, the plate and load (resistive) dissipation will be proportional to the respective areas shown on Fig. 9*b*. These areas are proportional to watts since their sides are represented by voltage and current having a product of voltamperes or watts. Thus the plate dissipation is $E_{bo} \times I_{bo}$ and the load dissipation is $(E_{bb} - E_{bo})I_{bo}$. With a sine-wave signal voltage applied to the grid the swing to point *h* on the load, as shown in Fig. 9*c*, will give an instantaneous plate dissipation represented by the area $A + 2B$.

$$\text{Plate dissipation at } h = A + 2B$$

A corresponding swing of the grid potential to point *l* in Fig. 9*d* shows a plate dissipation equal to $A + 2C$.

$$\text{Plate dissipation at } l = A + 2C$$

The grid signal voltage must pass through the *Q* point on each swing and at this point (quiescent) the plate dissipation is represented in Fig. 9*e* by the sum of the areas of the rectangles.

$$\text{Plate dissipation at } Q = A + B + C + D$$

The average value of the plate dissipation when the grid signal is applied must lie somewhere between the values of the extremes and its value at the *Q* point. It can be shown that this average value is the average of the areas at the extremes and *Q* point. Since the signal swings through the *Q* point twice on each cycle, the

Average value plate dissipation

$$= \frac{(A + 2B) + (A + 2C) + 2(A + B + C + D)}{4}$$

$$= A + B + C + \tfrac{1}{2}D$$

This average dissipation under operation is less than that for the quiescent or *Q* point by one-half the area of *D*.*

A study of Fig. 9*f* where the grid voltage swings from *h* to *l* and back to *l* shows that the total area in the large dotted rectangle is equal to $4D$ and corresponds to the numerator of equation 13. The cross-hatched area $\frac{1}{2}D$ is $\frac{1}{8}$ of $4D$ and is the average fundamental plate power as developed in equation 13. Thus the plate dissipation is reduced by the amount of the a-c power output.

* A mathematical solution shows that the reduction in plate dissipation varies as the sine squared. The average of the sine squared is one-half.

If the spacing between grid-curve intersections (Fig. 8) is not uniform, the output wave will not have the same shape as the input and amplitude or "harmonic" distortion will result. In the case of a typical triode the change in i_b will be greater as e_c swings toward the positive than for a corresponding negative change in e_c for operation in the negative grid region.

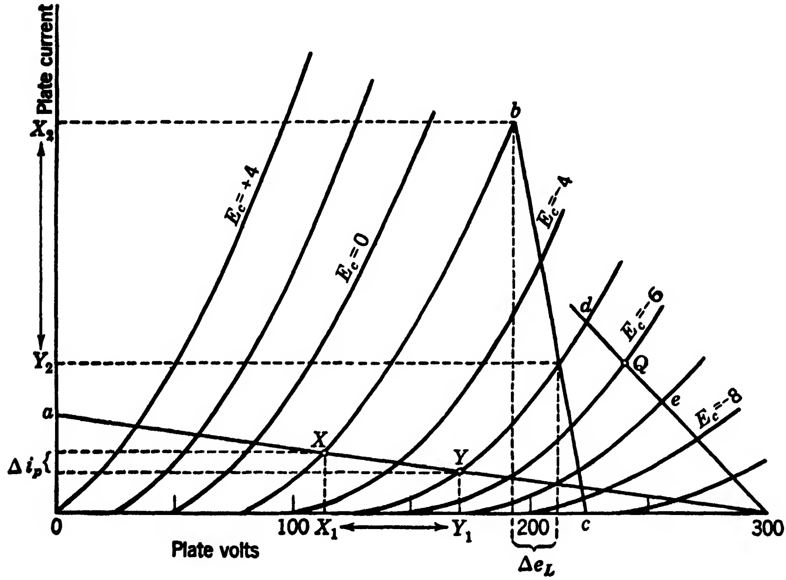


FIG. 10. Adjustments of the load line to secure best operating conditions.

The factors affecting the choice of the load for a vacuum tube are:

(*a*) The objective such as maximum gain, maximum voltage, maximum current, maximum power, or minimum distortion.

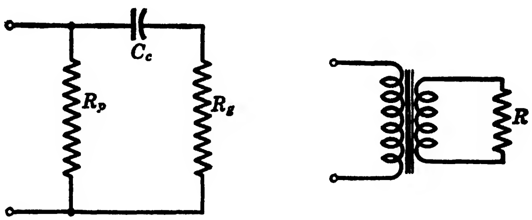
(*b*) Available power supply voltage.

(*c*) Maximum permissible plate dissipation and cathode current.

With these factors in mind, the best load line for any set of conditions may be found. For example, if maximum voltage gain is desired for the tube having the characteristics of Fig. 10 with a supply voltage of $E_{bb} = 300$, a load line through a and point 300 may be chosen. This load line represents a very high resistance load. For a grid voltage variation from X to Y the voltage variation across the load resistance will be X_1Y_1 with the corresponding plate current change of Δi_p. If a maximum current change is desired, a load line such as bc representing a low load resistance and a supply voltage of $E_{bb} = 225$ may be

chosen. For the same grid voltage shift (X to Y) this load line will produce a plate current change of X_2Y_2 and a voltage variation across the load of only Δe_L. As another example, if minimum distortion is desired, a load line such as de with a grid variation in the region of d to e will give the desired result. This follows because the spacing from d to Q and Q to e is equal.

Static and Dynamic Loads. In many cases the d-c and a-c resistances of the vacuum-tube load circuit are not equal. The d-c resistance is called the static load, and the a-c resistance is the dynamic load of the tube. Two load circuits are shown in Fig. 11 which have



(a) (b)

FIG. 11. Forms of plate-load circuits.

different static and dynamic loads. Part a of Fig. 11 has a higher d-c than a-c resistance and part b has a lower d-c resistance.

If in Fig. 11a the impedance of the B supply is negligible and the reactance of C_c is small compared to R_g, the a-c resistance is

$$R_{\text{a-c}} = \frac{R_pR_g}{R_p + R_g} \tag{15}$$

or simply the resistance of R_p in parallel with R_g. And

$$R_{\text{d-c}} = R_p \tag{16}$$

In Fig. 11b

$$R_{\text{a-c}} = \frac{R}{n^2} \tag{17}$$

where n is secondary to primary turns ratio of the transformer. The d-c resistance is merely the d-c resistance of the primary winding.

A convenient method for handling graphical solutions is to begin by plotting the static load line as described in the preceding article. The slope of this line is $-1/R_{\text{d-c}}$. With the slope known a load line may be constructed at any convenient place on the characteristic curves. Then a parallel line may be drawn through the operating

point by the use of triangles. The dynamic load line has a slope of $-1/R_{a\text{-}c}$, and the two load lines should intersect at a point corresponding to the average value of plate current and the value of grid bias selected (Fig. 12). As stated previously, if no distortion is present, the average value of plate current remains the same with and without signal. Hence, if distortion is negligible, the intersection of the static and the dynamic load lines coincides with the quiescent plate current which is defined by the intersection of the static load line and the grid-

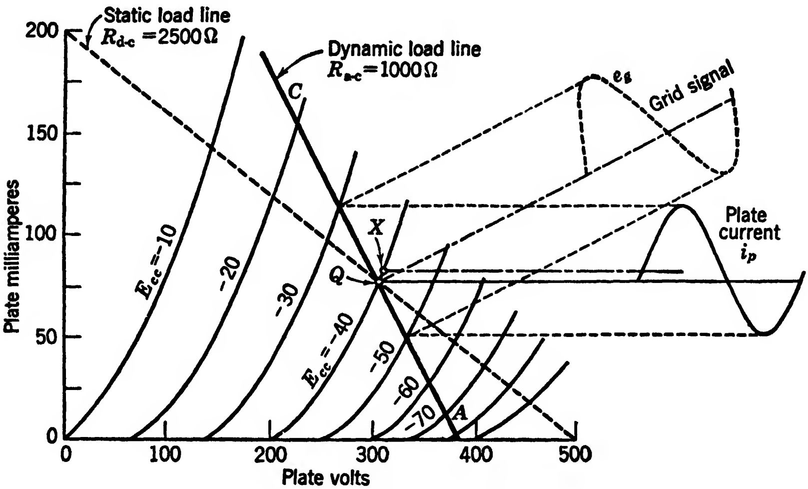


FIG. 12. Application of the dynamic load line to amplifier problems.

voltage curve corresponding to the grid bias used. Figure 12 illustrates the construction of the various load lines on the plate diagram of a vacuum tube. For this figure it is assumed that a plate supply of 500 volts, a d-c load resistance of 2500 ohms, an a-c load resistance of 1000 ohms for a given signal frequency, and a grid bias of -40 volt are to be used. The d-c load line will intercept the axis at 500 volts and 200 ma as shown by the dotted line. This load line crosses the -40-volt grid curve at point Q which gives the quiescent operating point. If a signal e_g is now impressed on the tube and the load resistance changes to 1000 ohms for $R_{a\text{-}c}$, a new a-c load line having a slope of $-1/R_{a\text{-}c}$ becomes effective. This load line AC is constructed so that it passes through the operating point Q. Now the impressed grid signal will swing along the a-c load line and cause the plate current to vary as shown on the right. The magnitude of the varying component of the plate current can be measured by horizontal projections to the Y axis. If amplitude distortion is not present, the posi-

tive and negative loops of the plate current will be equal and the operating point will remain at Q. In Fig. 12, the plate characteristic curves are so spaced that amplitude distortion does occur and the average value of the plate current is greater than the quiescent value. Accordingly, the actual operating point has moved up to point X on the $E_{cc} = -40$ curve. For this situation, if greater accuracy is desired, a new a-c load line should be constructed through point X parallel to the first line. Then the construction of grid signal and plate current curves should be repeated and a cut-and-try process is involved.

The a-c load resistance and the operating point should be chosen to represent the best compromise between the factors mentioned in the preceding article.

If the a-c load has a reactive component instead of a pure resistance, the dynamic load line will be an ellipse and as such it is not very useful. If the load is only slightly reactive, one may obtain satisfactory results using a straight line construction and assuming a resistance load equal to the impedance.

Effect of Interelectrode Capacitances. The interelectrode capacitances of a triode vacuum tube are indicated in Fig. 13. These capacitances are in a position to affect the input and output impedances of the circuits in which the tube is employed. The cathode-to-plate capacitance c_{pk} is in parallel with the load impedance, and the cathode-to-grid capacitance c_{gk} is in parallel with the input signal. The grid-to-plate capacitance c_{gp} may act to feed back energy from the output to the input and also may become very important as a factor in input impedance, as will be explained in the next paragraph. The magnitude of the interelectrode capacitances is of the order of a few micromicrofarads; hence their impedance is so large at low and medium frequencies that their presence can be neglected. At the higher audio frequencies and up they should be taken into account.

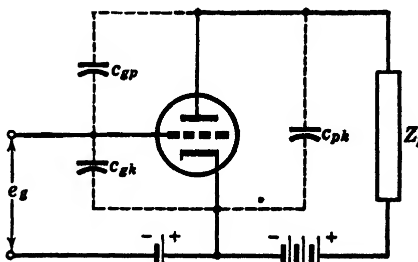


FIG. 13. Interelectrode capacitances in a triode circuit.

In multistage amplifiers operating at high frequencies, interelectrode capacitances become important as a part of output or load impedances and input impedances. The total capacitance load on the stage be-

tween two amplifier tubes using triodes is indicated in Fig. 14. In calculating these capacitances a phenomenon known as the Miller effect must be taken into account. Under suitable conditions the input capacitance to a tube is much greater when a load exists on the plate circuit than when it is on open circuit. This phenomenon can be visualized from Fig. 13 and the right section of Fig. 14. Thus, if a varying voltage e_g is impressed across the grid to cathode of the tube, the amplifying action will produce in the plate circuit a voltage of gain (G) times e_g. If the plate load is resistive, the resultant voltage across the grid to plate is $(e_g + Ge_g) = (1 + G)e_g$. The current

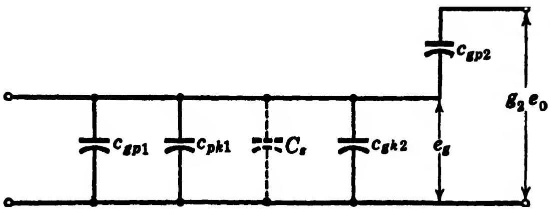


FIG. 14. Equivalent circuit illustrating the various capacitances in a two-stage amplifier.

which flows from grid to plate under this applied voltage is the same as would flow through a capacitance of magnitude $(1 + G)c_{gp}$ with an applied voltage of e_g. Thus the effective value of the grid-plate inter-electrode capacitance c_{gp} is greatly magnified by the Miller effect. The magnitude of the Miller effect is changed when the plate load is reactive and a phase angle with a resistive component is introduced. However, in most amplifier circuits the plate load arising from RC coupling or from a tuned tank circuit will be resistive so that the relationship derived above holds.*

The total capacitance units of Fig. 14 may be separated into (1) the output capacitance of the first tube, and (2) the input capacitance of the second tube. The output capacitance of the first tube is

$$C_0 = c_{gp_1} + c_{pk_1} + C_s \tag{18}$$

where C_s represents the stray capacity of the wiring from the tube to the coupling between the tubes. The magnitude of c_{gp_1} will be influenced in a very small measure by the Miller effect but this change may be neglected because calculations are based on plate voltage rather

* For more details concerning the Miller effect, see Chapter VII, *Radiotron Designer's Handbook*, 3rd ed., The Wireless Press, Sydney, Australia.

than grid voltage. In a similar manner, the input capacitance to the second tube for a resistive load will be

$$C_i = c_{gk_2} + (1 + G_2)c_{gp_2} \tag{19}$$

and the combined capacitance load C_t for the entire amplifier stage will be the sum of equations 18 and 19.

$$C_t = c_{gp_1} + c_{pk_1} + C_s + c_{pk_2} + (1 + G_2)c_{gp_2} \tag{20}$$

With screen-grid tubes and pentodes the grid-plate capacitance c_{gp} is very small and may be neglected at audio frequencies and low radio frequencies. With these tubes the input capacitance reduces to the c_{gk} plus the grid-screen capacity.

Problems

1. A power amplifier requires 5 watts of driving power to produce an output power of 60 watts. What is the power gain in decibels?

2. The power input to an amplifier is 2.5 microwatts and the output is 5 milliwatts. What is the loss or gain in db?

3. An amplifier has equal input and output resistances. If a 2-volt signal at the input produces a 37-volt signal at the output, what is the gain of the amplifier in db?

4. If the output resistance of the amplifier in Problem 3 is 600 ohms, express the power output level in vu (volume units).

5. A certain circuit contains a first amplifier with a gain of 12 db, a filter with a loss of 3 db, and a second amplifier with a gain of 6 db. What is the net gain or loss in the circuit? (15

6. The average power output of a certain crystal microphone is −48 db (reference level = 6 mw). This microphone is connected to an amplifier with a power gain of 80 db. What is the average power output of the amplifier?

7. A single-stage amplifier has a load resistance of 0.1 megohm. The vacuum tube has a plate resistance of 7700 ohms and an amplification factor of 20. What a-c plate current will flow under an a-c signal of 5 volts? What is the a-c voltage across the load resistance? What is the amplifier gain for this stage?

8. If the plate load for the amplifier circuit of Problem 7 is changed to an impedance having a resistance of 2000 ohms and an inductance L of 10 millihenries, what will be the gain at a frequency of 100,000 cycles?

9. In the amplifier shown in Fig. 15 $E_{bb} = 250$ volts; $E_{cc} = -3$ volts; $r_p = 25{,}000$ ohms; $\mu = 17.5$; $R_L = 150{,}000$ ohms; $I_b = 1.2$ ma; $E_g = 2.0$ volts (rms). (a) Calculate I_p. (b) Calculate E_p. (c) What is the voltage gain of the amplifier?

10. The amplifier in Problem 9 is reconnected as shown in Fig. 16. The voltage drop across R_k provides bias for the tube. (a) Calculate the proper

value for R_k. (b) What would be a suitable value for C_k if the lowest frequency to be amplified is 75 cps? Note: C_k is usually designed to have a reactance in ohms equal to 0.1 the resistance of R_k, also in ohms, at the lowest frequency to be amplified.

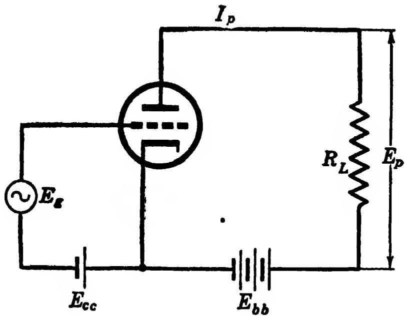


FIG. 15

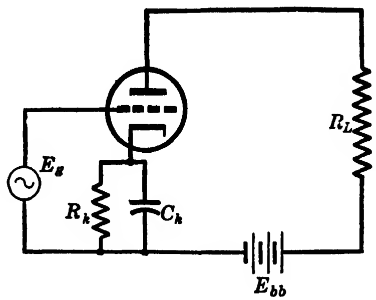


FIG. 16

11. Draw the equivalent circuit of the amplifier shown in Fig. 17. E_{bb} = 250 volts; $E_{cc} = -3$ volts; $\mu = 17$; $E_g = 2$ volts; $I_b = 0.8$ ma; $r_p = 40{,}000$; $R_p = 250{,}000$ ohms; $R_g = 500{,}000$ ohms; $C_c = 0.005$ μf. (a) Calculate E_0 when the frequency is 40 cps. (b) Calculate E_0 when the frequency is 5000 cps.

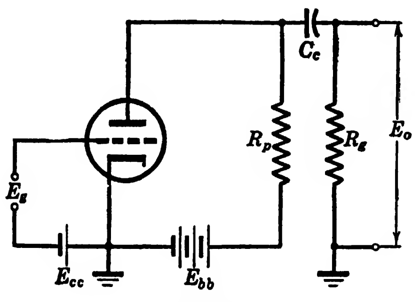


FIG. 17

12. A 6J5 tube is connected as shown in Fig. 17. R_p = 150,000 ohms; R_g = 300,000; E_g = 2.12 volts; E_{bb} = 300 volts; $E_{cc} = -3$ volts. (a) Use the plate family of characteristics (page 63) and draw the static load line. What is the value of I_{bo}? What power must be dissipated by the plate of the tube? (b) Assume that C_c is large enough to be neglected and determine the dynamic load line. Calculate E_0 and the voltage gain. (c) If cathode bias were used instead of fixed bias, what value should the cathode bias resistor have?

13. The amplifier in Fig. 18 uses a 6L6 (page 87) connected as a triode. E_{bb} = 275 volts; $E_{cc} = -22.5$ volts; E_g = 15.9 volts; resistance of loud-speaker voice coil = 8 ohms; primary to secondary turns ratio = 25 to 1; d-c

resistance at primary = 375 ohms. (a) Draw the static load line. Find I_{bo} and the plate dissipation. (b) Draw the dynamic load line. Calculate I_p and E_p. (c) What is the power output of the amplifier? (d) What is the value of $E_{b\ max}$? Why is $E_{b\ max}$ greater than E_{bb}?

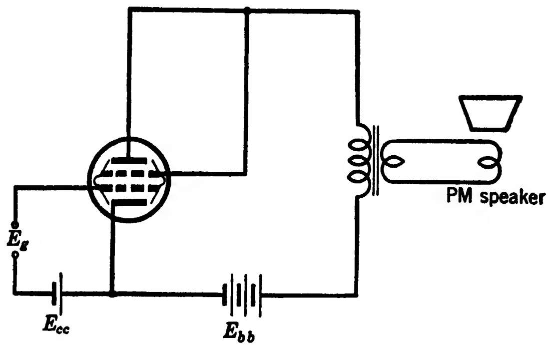


FIG. 18

14. In the amplifier shown in Fig. 19, $\mu = 18$; $r_p = 30{,}000$ ohms; $R_L = 220{,}000$ ohms; $C_s = 12.3$ $\mu\mu$f; $c_{gp} = 3.4$ $\mu\mu$f; $c_{gk} = 3.4$ $\mu\mu$f; $c_{pk} = 3.6$ $\mu\mu$f. What is the input capacitance of the amplifier?

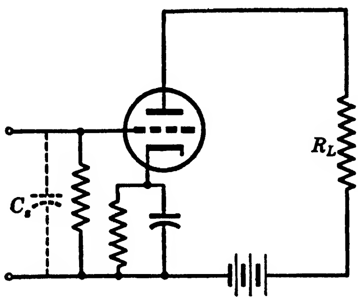


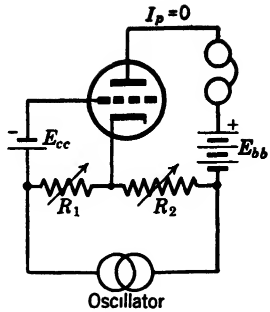
FIG. 19

15. A single-stage amplifier uses a tube having a transconductance of 2600 micromhos and a plate resistance of 8000 ohms. If the plate load consists of a 5000-ohm and a 3000-ohm resistor in parallel, what is the gain of the amplifier circuit? If an a-c signal of 8 volts is applied to the input, what plate current will flow? What will be the a-c voltage across the load resistors?

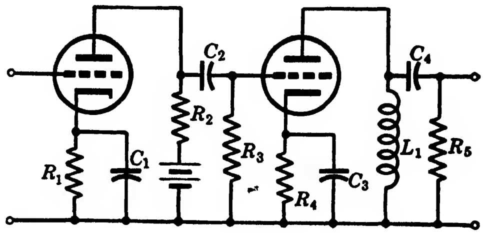
16. A certain pentode has a transconductance of 2000 micromhos and a plate resistance of 800,000 ohms. What will be the gain of a single-stage amplifier using this tube with a load resistance of 20,000 ohms?

17. Miller bridge for determining μ of a triode. R_1 and R_2 are adjusted until zero current flows in the headphones. Draw the equivalent circuit and show that

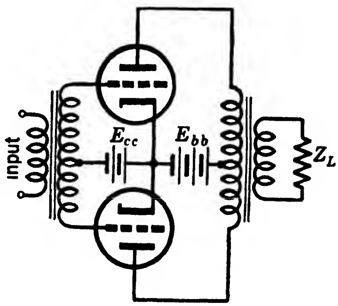
$$\mu = \frac{R_2}{R_1}$$



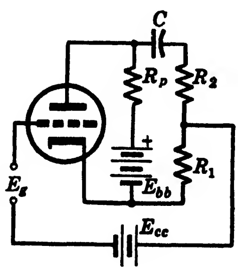
18. Draw equivalent circuit.



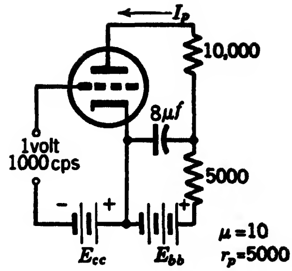
19. Draw equivalent circuit.



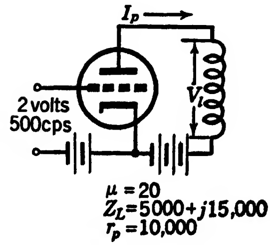
20. Draw equivalent circuit.



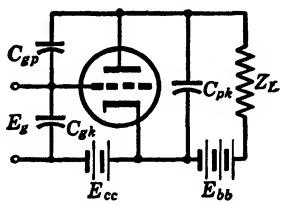
21. Draw equivalent circuit. Find I



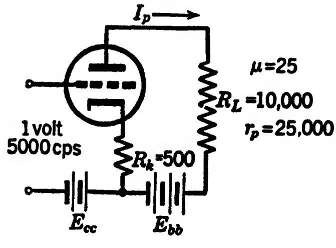
22 Draw vector diagram. Find I_p, V_l, and voltage gain. How much does output voltage lag input voltage?



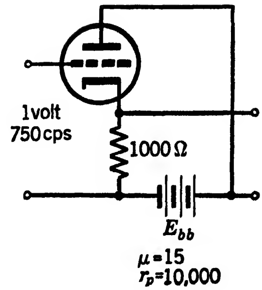
23. Draw equivalent circuit.



24. Draw equivalent circuit. Find I_p and voltage gain.



25. Draw equivalent circuit and find voltage gain.



Chapter VII

MULTISTAGE VOLTAGE AMPLIFIERS

Multistage Amplifiers. The one-tube, single-stage amplifier discussed in the preceding chapter frequently fails to produce as much gain as is desired. Additional gain may be secured by coupling two or more amplifier stages in *cascade* wherein the output voltage of one stage is fed into the grid input of the succeeding stage. In such multistage operation the over-all voltage gain will be the product of the voltage gains of the individual stages. The over-all gain may be carried to a magnitude where *instability* of operation will result, or where the circuit noise and microphonics introduced will become objectionable. Hence care must be exercised in the design of cascade amplifying units. The type of coupling that is chosen for the connection between the individual stages is important and depends upon a number of factors, the chief of which is the nature of the signal frequency being amplified. Multistage or cascade amplifiers may be classified as follows:

(*a*) *Direct-coupled amplifiers*—primarily those which operate at 0 to 15,000 cycles per second.
(*b*) *Audio amplifiers*—covering 20 to 15,000 cycles per second.
(*c*) *Video amplifiers*—covering 20 to 5,000,000 (or higher) cycles per second with *uniform phase delay and frequency response.*
(*d*) *Radio-frequency amplifiers*—operating in the region above the audible range and ordinarily capable of passing only a relatively narrow band of frequencies.

Direct-Coupled Amplifiers. For applications such as electrocardiography, the study of nerve currents, and the use of an oscilloscope to show transient responses relative to a steady component, it is necessary to use a circuit that will amplify d-c voltages as well as a-c voltages. Circuits for these applications are often called direct-current (d-c) amplifiers. A circuit for a direct-coupled or d-c amplifier is shown in Fig. 1. Two practical limitations prevent the use of this circuit except where absolutely necessary. These limitations are: (1) Each stage of the amplifier must be provided with a separate source

of power supply (battery or power pack), and (2) the magnitude of the "drift" in the operating point of the final stage due to changes in the voltages or circuit constants of the preceding stages. Even a very minute change in the plate potential of the first stage, when amplified by the following stages, may be sufficient to shift the final grid far beyond its operating range. This becomes so serious in a high-gain amplifier that more than two stages are seldom used without special precautions such as supply voltage regulation, balanced circuits, or degenerative feedback.

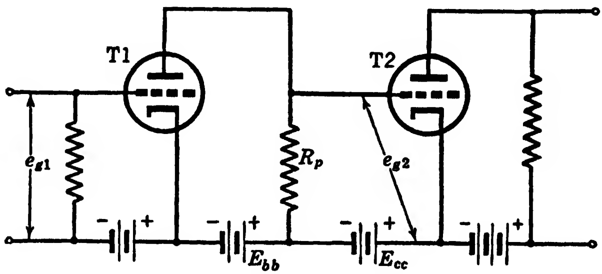


FIG. 1. Direct coupling of amplifier stages.

At low and medium frequencies the gain per stage may be calculated directly from equations 6 or 10 of Chapter VI. At high frequencies where the reactance of the shunting capacities becomes appreciable, the information given by equations 18, 19, and 20 of the preceding chapter should be considered.

Audio Amplifiers. Resistance-Capacity Coupling. The difficulties of the direct-coupled circuit are overcome by using a coupling condenser C_c between the amplifier stages as shown in Fig. 2. This coupling condenser C_c prevents the d-c plate-supply voltage of tube 1 from affecting the grid of tube 2, but permits the alternating component of voltage drop across R_p to be applied to R_g and the cathode-grid input of tube 2. It will be noted that the output voltage of tube 1 across R_p (e_z) is applied across C_c and R_g in series, acting as a potential divider. Hence the a-c input to the grid of tube 2 is that part of the a-c component across the grid-leak resistor R_g. The resistance of R_g is large, up to about 1 megohm, while the reactance of C_c will vary with the signal frequency. Thus the lowest frequency that can be amplified is a function of R_g and the reactance of the coupling condenser X_c. Since $X_c = 1/2\pi f C_c = \infty$ at $f = 0$, the amplifier will not amplify direct current. The bias on grid 2 is determined by the battery E_{cc} and is not upset by slow "drifts" in the plate voltage of the pre-

ceding tube so that common power supplies can be used. This type of coupling is commonly designated as an RC coupling.

The operation of the RC coupling circuit can be analyzed conveniently by applying the equivalent circuit theory to three ranges of

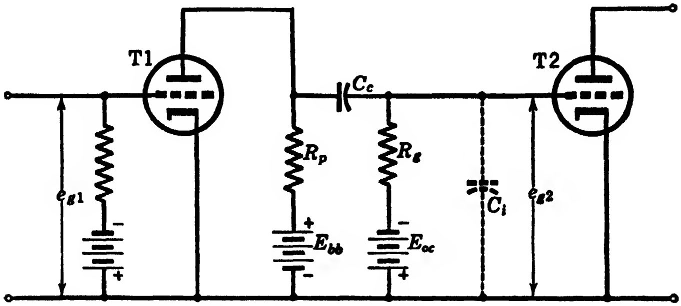


FIG. 2. Resistance-capacity coupling of amplifier stages.

frequency—middle, low, and high. Either the constant-voltage or the constant-current equivalent circuit may be used, though the latter serves better for the parallel type of load which exists in multistage amplifiers. An approximate equivalent plate load circuit for a middle

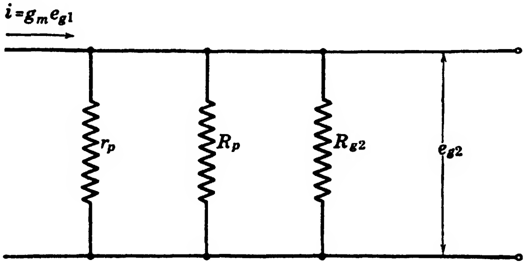


FIG. 3

frequency range is shown in Fig. 3. The reactance of C_c is negligible and has been omitted. To simplify the gain equations for purposes of comparison, the following new symbols will be adopted.

Let

$$R_S = r_p \text{ and } R_p \text{ in parallel} = \frac{r_p R_p}{r_p + R_p} \tag{1}$$

and

$$R_L = r_p,\ R_p, \text{ and } R_{g2} \text{ in parallel} = \frac{r_p R_p R_{g2}}{r_p R_p + r_p R_{g2} + R_p R_{g2}} \tag{2}$$

Then, since all resistors are in parallel, e_{g2} is the same as the drop across R_p, and the gain from equation 11, Chapter VI, is

$$G_{\text{middle frequencies}} = -g_m R_L \tag{3}$$

$$= -\frac{g_m r_p R_p R_{g2}}{r_p R_p + r_p R_{g2} + R_p R_{g2}} \tag{4}$$

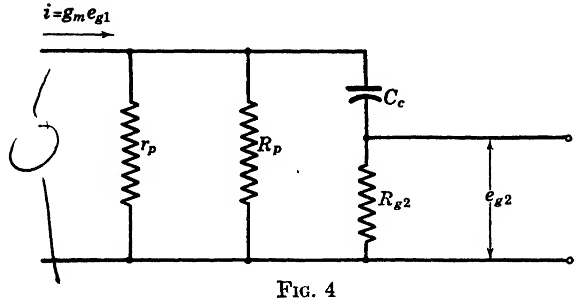


Fig. 4

For the low-frequency range (Fig. 4) the coupling capacitor C_c lowers the amplified gain across R_p in the voltage divider ratio $R_{g2}/(R_{g2} - jX_c)$. The gain across R_p is

$$G = -g_m \frac{R_S(R_{g2} - jX_c)}{R_S + R_{g2} - jX_c}$$

Hence the gain across R_{g2} is the product of the voltage divider ratio and the gain.

$$G_{\text{low frequencies}} = -\frac{g_m R_S R_{g2}}{R_S + R_{g2} - jX_c} \tag{5}$$

$$|G| = \frac{g_m R_S R_{g2}}{\sqrt{(R_S + R_{g2})^2 + X_c^2}} \tag{6}$$

The approximate equivalent load circuit for the high-frequency range is shown in Fig. 5. Here the reactance of the coupling condenser C_c is negligible but the shunt capacitance C_t which is the equivalent sum of the tube interelectrode capacitances and the wiring becomes effective. The total plate load is that of R_L and the reactance of C_t in parallel, and the gain is

$$G_{\text{high frequencies}} = -\frac{g_m R_L(-jX_{C_t})}{R_L - jX_{C_t}} \tag{7}$$

$$|G| = \frac{g_m R_L X_{C_t}}{\sqrt{R_L^2 + X_{C_t}^2}} \tag{8}$$

The formula for C_t was developed in the preceding chapter (equation 20).

$$C_t = c_{gp1} + c_{pk1} + C_s + c_{gk2} + (1 + G_2)c_{gp2}$$

The last term of this equation is important in high-μ triode amplifiers as shown by the following example. A typical high-μ triode has r_p = 66,000, μ = 100, c_{gp} = 2.4, c_{gk} = 4, and c_{pk} = 3.6 $\mu\mu$f. If a value of 15 $\mu\mu$f is allowed for C_s, then the sum of the first three terms of the above equation is 15 + 3.6 + 4 = 22.6 $\mu\mu$f. Assuming a gain of 0.6μ, the last term will be (1 + 60)2.4 = 146.4 $\mu\mu$f. The total C_t will be 22.6 + 146.4 = 167 $\mu\mu$f, which at 15,000 cycles has a reactance of 63,600 ohms.

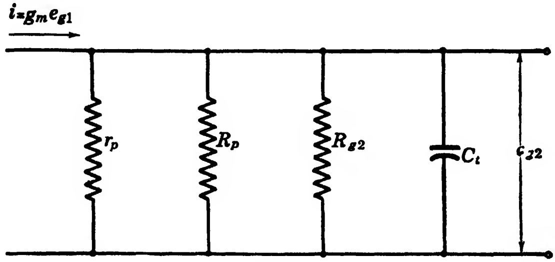


FIG. 5

The c_{gp} of any tube containing a screen grid is negligible, being for a typical pentode only 0.005 $\mu\mu$f. This causes the last term of the equation to be negligible and the resulting C_t to be small. For this reason screen-grid tubes are preferable in high-gain voltage amplifiers whenever constant gain at the higher frequencies is desired.

A comparison of the gain at low or high to the middle frequency range can be made by calculating the ratio of the respective gains from the preceding equations.

$$\text{Gain}\,\frac{\text{low}}{\text{middle}} = \frac{\dfrac{-g_mR_SR_{g2}}{R_S + R_{g2} - jX_c}}{-g_mR_L} = \frac{\dfrac{R_SR_{g2}}{R_S + R_{g2} - jX_c}}{\dfrac{R_SR_{g2}}{R_S + R_{g2}}}$$

$$= \frac{R_S + R_{g2}}{R_S + R_{g2} - jX_c} = \frac{1}{1 - \dfrac{jX_c}{R_S + R_{g2}}}$$

$$|G|\,\frac{\text{low}}{\text{middle}} = \frac{1}{\sqrt{1 + \dfrac{X_c^2}{(R_S + R_{g2})^2}}} \tag{9}$$

Thus the amount that the gain falls off at low frequencies is determined by the ratio of the reactance X_c of the coupling capacitor to the sum of resistance of the grid resistor R_g in series with R_S (R_L and r_p in parallel). A convenient reference point for a comparison exists where $X_c = R_S + R_{g2}$. Here the low-frequency gain is 70.7 per cent of the middle frequency gain. This is the "half-power" point, or, expressed in decibels, the frequency at which the amplifier is "down 3 db." Since the designer has more interest in gain with respect to frequency than in the impedance ratio, the frequency f_0 for the half-power point should be calculated.

$$\frac{1}{2\pi f_0 C_c} = R_S + R_{g2}$$

$$f_0 = \frac{1}{2\pi C_c(R_S + R_g)} \tag{10}$$

The comparative gain at low frequencies in terms of a ratio to f_0 has been plotted in Fig. 6 and may be used conveniently in solving problems.

The ratio of the gain at high frequencies to middle frequencies is

$$\text{Gain}\,\frac{\text{high}}{\text{middle}} = \frac{\dfrac{g_m R_L X_{C_t}}{\sqrt{R_L^2 + X_{C_t}^2}}}{g_m R_L} = \frac{X_{C_t}}{\sqrt{R_L^2 + X_{C_t}^2}}$$

$$= \frac{1}{\sqrt{1 + \left(\dfrac{R_L}{X_{C_t}}\right)^2}} \tag{11}$$

Thus the amount that the gain falls off at high frequencies is determined by the ratio of the total shunt resistance R_L to the shunt capacitive reactance X_{C_t}. Again, a convenient reference point exists where $R_L = X_{C_t}$, which gives a 70.7 per cent gain and half-power point. The frequency f_0' at which this occurs makes a good reference value. The gains for high frequencies in terms of a ratio to f_0' are plotted in Fig. 6 for use in solving problems.

The effect upon the amplifier gain of varying the magnitude of coupling capacitor C_c is illustrated in Fig. 7. A corresponding effect upon gain at high frequencies arising from a variation of the shunt capacitance C_s or C_t is shown in Fig. 8.

The preceding equations for gain at low and high frequencies reveal two kinds of distortion different from the harmonic or amplitude type.

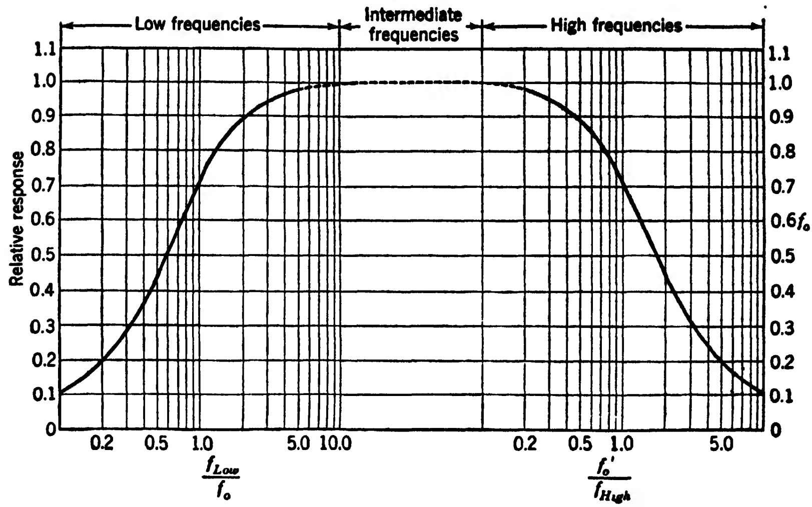


FIG. 6. Relative response of resistance-coupled amplifiers at low and high frequencies (f_0 and f_0' are frequencies for the half-power point).

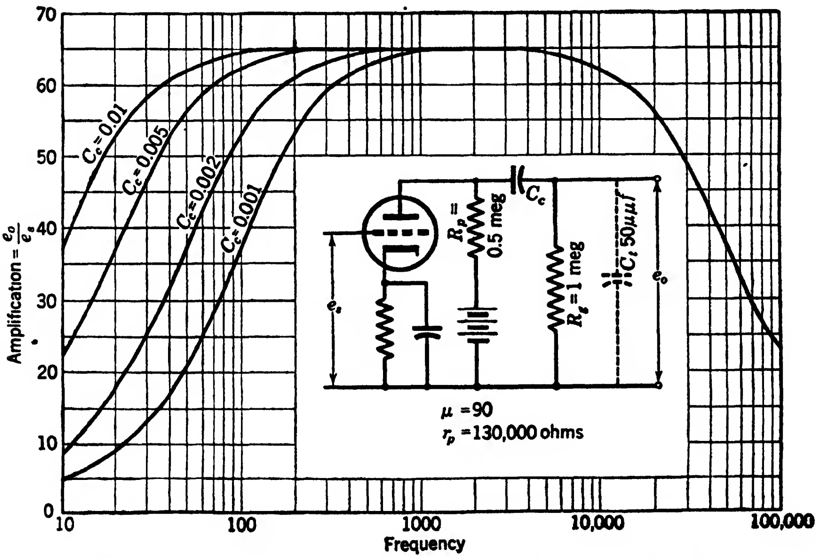


FIG. 7. Effect of magnitude of capacitance coupling on frequency response of amplifiers.

These are frequency distortion wherein gain varies with frequency, and phase distortion wherein e_{g2} does not remain at an angle of -180 degrees with respect to e_{g1}.

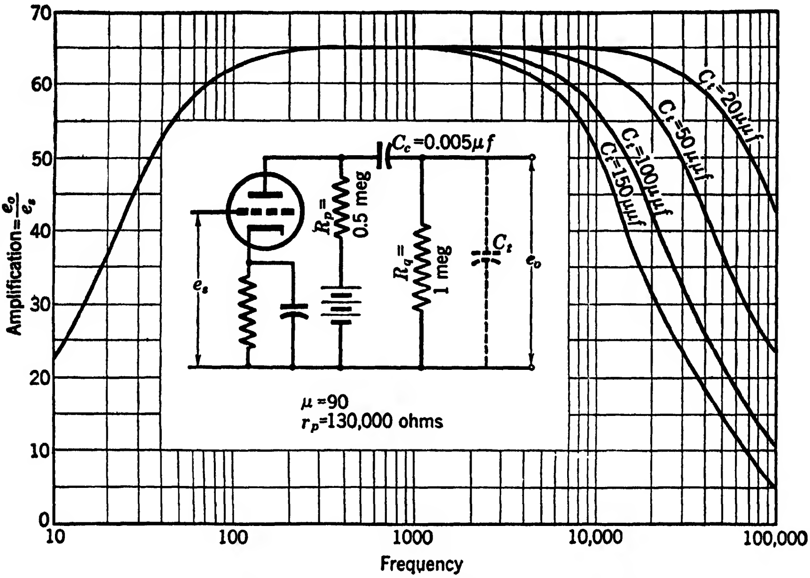


FIG. 8. Effect of magnitude of shunt capacitance on magnitude of frequency response.

Audio Amplifiers. Impedance Coupling. It is sometimes desirable to use inductances instead of employing resistors for coupling elements. A single inductance may be used to replace either the grid or the plate resistor, or both resistors may be replaced as in Fig. 9. The advantage of using inductance in the plate circuit is that the d-c voltage drop will be negligible because of the comparatively low d-c resistance of the winding. This low voltage drop permits the operating point to be at almost full E_{bb} supply potential with a corresponding increase in maximum plate swing (E_{pm}). Such a saving in effective supply voltage is particularly desirable in battery-operated portable

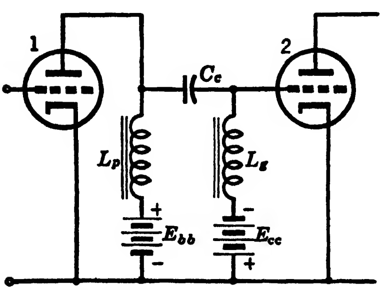


FIG. 9. Inductance coupling of amplifier stages.

equipment, although the advantage is canceled somewhat by the weight and bulk of the coupling inductances.

Inductances are often useful in the grid circuit of power tubes having characteristics that demand a low resistance path from grid to cathode. However, for most applications impedance coupling offers no advantages over transformer coupling and is usually more expensive.

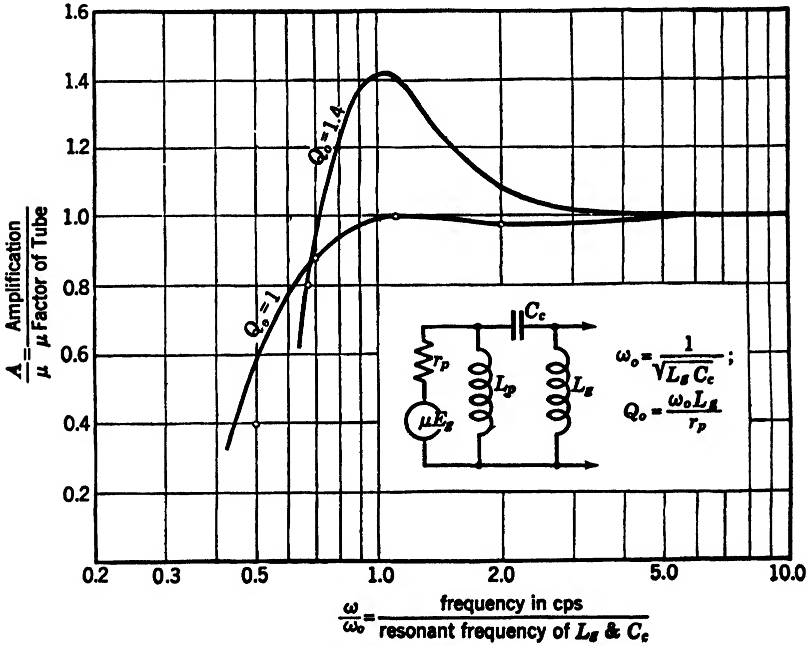


FIG. 10. Low-frequency response of an inductance-coupled amplifier.

Whenever the *effective resistance* of the coils is negligible compared to their reactance and the impedance of the coupling condenser C_c is low, gain can be stated simply as

$$G = -\frac{\mu j X_L}{r_p + jX_L} \tag{12}$$

where X_L is the reactance of L_p and L_g in parallel (Fig. 9). Since $X_L = 2\pi f L$, gain will increase with frequency until a maximum of μ is reached.

In practice, core and copper losses and saturation of the core at low frequencies reduce the effective impedance of the coils. However, at the low frequency where $X_c = X_{Lg}$ series resonance occurs and the

gain may actually exceed μ. For these reasons equation 12 must be used with care and a more nearly exact expression applied if the problem demands. A typical gain-frequency curve for an impedance-coupled amplifier is given in Fig. 10.

Audio Amplifiers. Transformer Coupling. The grid and plate coils of an amplifier may be placed on the same core to form a transformer as shown in Fig. 11. This eliminates the necessity for a coupling condenser and permits an increase in voltage gain through the use of a greater number of turns on the grid winding than on the plate. For an ideal transformer having zero losses and no leakage the relationship between primary and secondary is

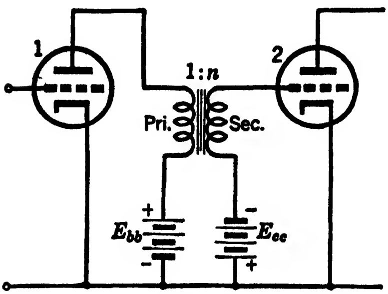


FIG. 11. Impedance and transformer couplings for amplifiers.

$$\frac{E_1}{E_2} = \frac{I_2}{I_1} = \frac{N_1}{N_2} = n \tag{13}$$

$$\frac{Z_1}{Z_2} = \left(\frac{N_1}{N_2}\right)^2 = n^2 \tag{14}$$

where 1 denotes the primary, 2 the secondary, N the number of turns, and n the ratio of transformation. Z_1 is the impedance looking into the primary with a load Z_2 connected across the secondary.

For the ideal case gain can be easily found if it is realized that the plate load is simply an impedance X_{L1} (the reactance of the transformer primary) and that the voltage at grid 2 is n times the voltage at plate 1.

$$G = \pm \frac{\mu j X_{L1}}{r_p + jX_{L1}} n \tag{15}$$

$$|G| = \frac{\mu X_{L1}}{\sqrt{r_p^2 + X_L^2}} n \tag{16}$$

The signal voltage e_{g2} may be either positive or negative with respect to e_{g1}, depending on the manner in which the secondary is connected relative to the primary. This feature is advantageous where it is necessary to reverse the phase of a signal without adding another amplifier stage.

The practical transformer is subject to all the difficulties encountered with coupling impedances plus several more of its own as follows:

(a) Core loss due to hysteresis and eddy currents.

(*b*) Core saturation at low frequencies (aggravated by the d-c or steady component of plate current).

(*c*) Copper loss due to the effective resistance of the windings.

(*d*) Leakage reactances caused by failure of the flux produced by one winding to link all the turns of both windings.

(*e*) Distributed capacities between turns and between windings.

(*f*) Hum pick-up from stray magnetic fields.

In designing a transformer to have minimum core loss careful consideration must be given to the iron used in the core. Eddy currents are reduced by the use of thin laminations. In practice these run from 0.015 inch for the ordinary class of transformers to as thin as 0.001 inch where extremely high frequencies are encountered. In order to obtain maximum impedance in the primary winding, the core should be tightly stacked and should be of sufficient cross section so the iron always operates below the knee of its magnetization curve. If the circuit is such that there is a d-c component of current in one winding, the point of operation will be pushed upward on the curve and the flux due to the a-c component may be distorted if it swings across the knee of the curve. This effect is apparent especially at low frequencies and gives rise to a large third harmonic of the fundamental frequency. The best way to prevent this saturation is to eliminate d-c currents from the windings. This may be done by using a resistance or an impedance coil between E_{bb} and the plate of the tube (Fig. 12). The transformer primary is then connected to the tube through a coupling condenser. Another method is to introduce an air gap in the magnetic circuit of the transformer. This reduces the flux density in the iron and lowers the primary inductance. However, the *incremental primary inductance* (the inductance to varying currents) will increase with the length of the air gap to a maximum before dropping off.*

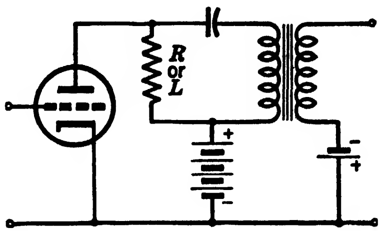


FIG. 12. Circuit for elimination of direct current in a transformer-coupled amplifier.

Copper loss can be minimized by using a wire of large cross section, few turns, and a short mean length per turn, but this practice is inconsistent with high inductance and large turns ratio. Also, at high frequencies the effective resistance of large wire increases due to the skin

* *Principles of Radio Engineering,* Glasgow, McGraw-Hill, 1936, pp. 65–67.

effect. If the windings are made with alternate layers of primary and secondary, flux leakage will be low. However, this practice may cause the interwinding capacitance to become excessive. Few turns and a grounded shield between primary and secondary will provide the smallest capacitance coupling.

If the transformer is to be used in a low-level circuit, special attention must be given the problem of hum pick-up. Stray magnetic fields

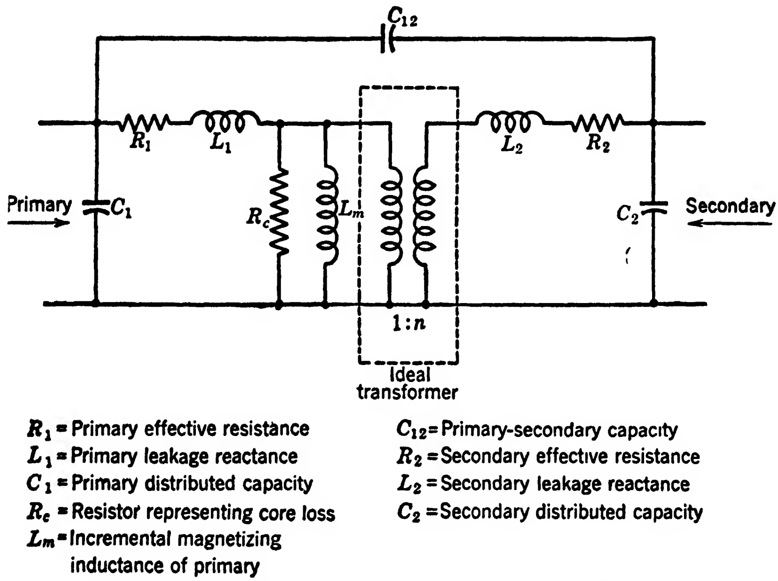
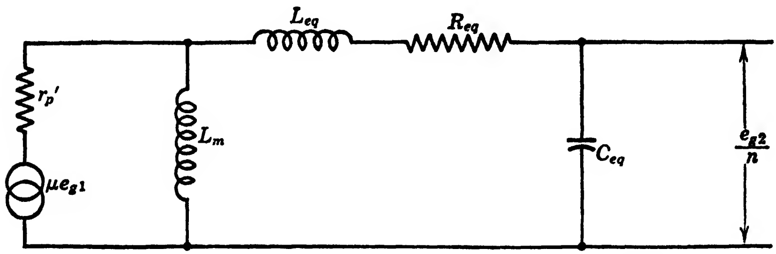


FIG. 13. Complete equivalent circuit of a transformer.

from the power supply components or from nearby power circuits may induce currents in the transformer many times that of the desired signal. In the better grade of transformers two sets of windings are placed on the core in such a manner that the hum voltages induced in one are exactly canceled by those induced in the other. Some benefit can also be obtained by enclosing the core and coil assembly in a thick iron shield having tightly fitted joints.

The complete equivalent circuit of a practical transformer is somewhat involved but the input impedance, voltage ratio, and phase shift may be calculated from the circuit theory given in some textbooks on a-c machinery. A complete equivalent circuit is given in Fig. 13. If some of the terms in the complete circuit which normally are not very important, are omitted and others combined, a fairly simple circuit (Fig. 14) may be evolved which can be used for most applications.

This circuit, in turn, may be resolved into simple equivalent circuits for calculating gain at the middle, low, and high frequencies, as illustrated in Fig. 15. At the middle frequencies where the reactances of



$r_p' = r_p + R_1$

$L_{eq} = L_1 + \frac{L_2}{n^2}$

$R_{eq} = \frac{R_2}{n^2}$

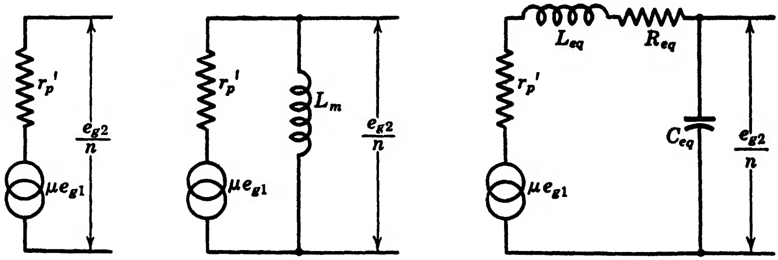
$C_{eq} = n^2(C_2 + C_i + \frac{n \pm 1}{n} C_{12})$

C_1, C_{pk}, R_g, R_c omitted

Note: Sign in term $\frac{n \pm 1}{n}$ depends on connection of secondary relative to primary.

FIG. 14. Simplified equivalent circuit of a transformer and tube; secondary terms refer to primary.

L_m and C_{eq} are high compared to r_p', gain is $\pm\mu n$. At low frequencies gain is reduced by the reactance of L_m, and at high frequencies by the capacity C_{eq}. The solution of the equations given in Fig. 15 will give both gain and phase shift.



$G = \pm\mu n$

$G = \pm \frac{jx_{Lm}}{r_p' + jx_{Lm}} \mu n$

$G = \mp \frac{jx_{C_{eq}}}{r_p' + R_{eq} + jx_{L_{eq}} - jx_{C_{eq}}} \mu n$

(a) Middle frequencies (b) Low frequencies (c) High frequencies

FIG. 15. Simplified equivalent circuits for transformer-coupled amplifiers.

The peak near the high end of the response curve, Fig. 16, is caused by the series resonance of the equivalent leakage reactance L_{eq} and the equivalent shunt capacity C_{eq}. Since C_{eq} contains a term

$[(n \pm 1)/n]C_{12}$, there will be two values of capacity, depending on whether the secondary is connected "capacity aiding" or "capacity bucking." The height of the peak depends on the effective Q of the circuit and may be lowered by increasing the resistance of the secondary winding or by using a load resistor across the secondary.

The plate resistance of tube 1 and the input capacity of tube 2 are important factors in the gain equations. Hence it is necessary for the transformer to be designed specifically for the tubes with which it is to be used. Tubes containing screen grids, because of their high plate resistance, show wide variations in gain with frequency when they are

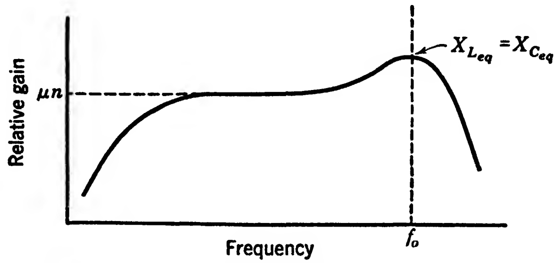


FIG. 16. Response of typical transformer-coupled amplifier stage.

transformer coupled and are therefore used only when the transformer is loaded, as in a power output stage.

Video Amplifiers. The term video frequency comes from the television art and is applied to those frequencies which go to make up a television picture. These frequencies extend from the low audio through the supersonic and radio range to 5 megacycles or higher. A video amplifier is one that will pass this range of frequencies with substantially constant amplitude and with uniform time delay. The phase relationship of the various component frequencies in a television picture is extremely important, and usually more attention is devoted to obtaining linear phase shift than to securing uniform amplitude response or freedom from harmonic distortion. The importance of phase shift is illustrated in Fig. 17 where an approximate square wave has been created by superimposing a third harmonic of suitable amplitude on a fundamental. A more nearly square wave form would be produced by the addition of the fifth harmonic. The effect on the wave form of a phase angle shift of the third harmonic is illustrated by the resultant wave at the right of Fig. 17. Television signals are frequently of the square-wave type.

The major differences betweer a video amplifier and an audio amplifier are in the high-frequency region. The response of a resistance-

coupled amplifier drops at high frequencies because of the shunting capacity of the tube and circuit elements. Also the grid-plate capacity is extremely important as shown by equation 20, Chapter VI. In order to keep this capacitance at the minimum, pentode tubes are almost always used for video amplifiers. The effect of the various shunting capacities is reduced by a plate load of relatively low resistance. This results in low gain per stage and has led to the development of pentode tubes having extremely high mutual conductance such as the 6AC7, 6AG7, and others.

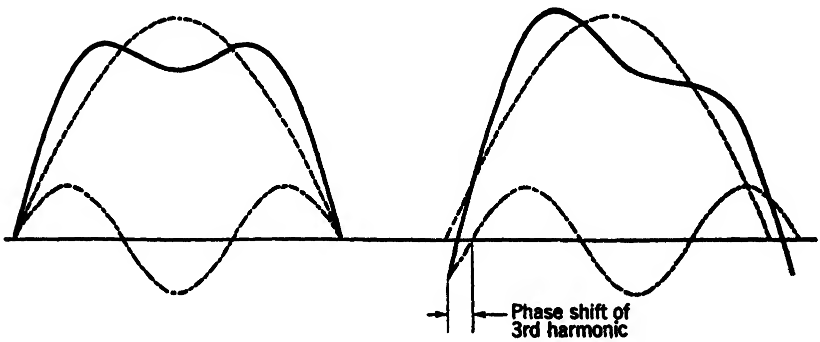
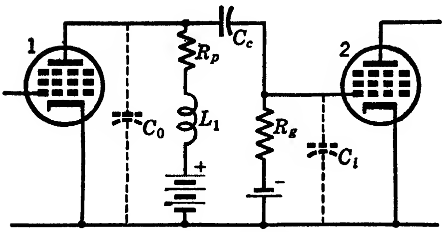


FIG. 17. Effect of a small phase shift of a harmonic on the wave form of a square type of wave.

Further improvement of the high-frequency response is provided by small inductance coils connected in combination with the plate-load resistance. The simplest of these circuits is the "shunt-peaked" combination of Fig. 18*a*, where an inductance is placed in series with the plate-load resistance. The grid resistance R_g is made high (consult tube manual), and the coupling condenser C_c has a low reactance at the lowest frequency to be passed. C_t represents the total shunting capacity which is the sum of the output capacity of tube 1, the input capacity of tube 2, and the stray wiring capacities in the circuit. For a satisfactory compromise between uniform gain and constant phase shift, the load resistance R_p should be equal to the reactance of C_t at the highest frequency (f_0) that is to be passed. The reactance of the peaking coil L_1 is equal to half the reactance of C_t at f_0. Since R_p is very low compared to R_g and the internal plate resistance of tube 1, stage gain can be easily found by multiplying the mutual conductance G_m by R_p.

In the shunt-peaked circuit, it was necessary to add the output capacitance of stage 1 and the input capacitance of stage 2 when



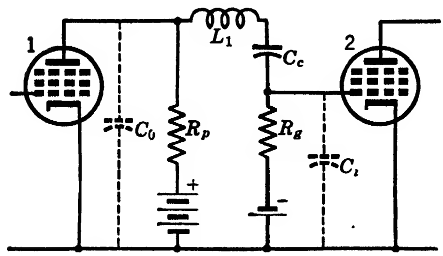
$C_t = C_0 + C_i$

$R_p = \frac{1}{2\pi f_0 C_t} = X_{Ct}$ at f_0

$L_1 = \frac{R_p}{4\pi f_0} = \frac{X_{Ct}}{4\pi f_0}$

$G = -g_m R_p$

(*a*) Shunt



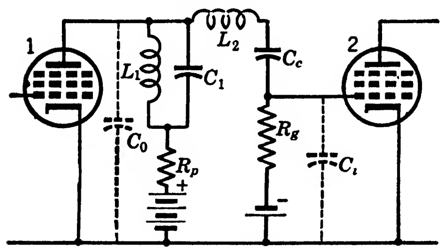
$C_i = 2C_0$

$C_t = C_0 + C_i$

$R_p = 1.5X_{Ct}$ at f_0

$L_1 = \frac{3X_{Ct}}{4\pi f_0}$

(*b*) Series



$m = 0.6$

$C_0 = 0.8C_i$

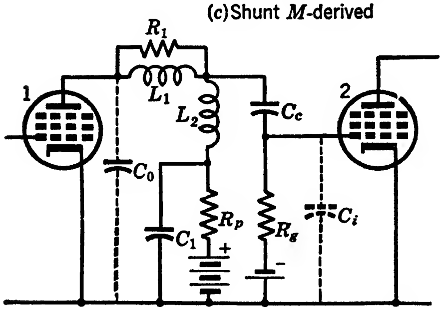
$C_1 = 0.533\,C_i$

$R_p = 2X_{Ci}$ at f_0

$L_2 = \frac{R_p}{\pi f_0}$

$L_1 = \frac{mL_2}{2} = 0.3L_2$

(*c*) Shunt *M*-derived



$C_i > 2C_0$ or $2C_1$

$R_p = 2X_{Ci}$ at f_0

R_1 by experiment

$L_1 = (1/2 + \frac{C_0}{C_i})\frac{R_p}{\pi f_0}$

$L_2 = (1/2 + \frac{C_1}{C_i})\frac{R_p}{\pi f_0}$

(*d*) 2-Section constant-*K*

FIG. 18

calculating the load resistance. A more desirable arrangement is the "series-peaked" circuit wherein a coil L_1 is placed between plate and grid (Fig. 18*b*). Here the plate and grid capacities are separated by the inductance, and the circuit assumes the characteristics of a π coupling network. The load resistance in this case is $\frac{3}{2}$, the reactance of the total shunting capacity, and provides 50 per cent more gain than the shunt-peaked circuit. However, for proper operation of this circuit C_t must equal twice the capacity of C_0. This is approximately true with the usual pentodes but may be made exactly true in all circuits by the addition of a small trimmer condenser. The

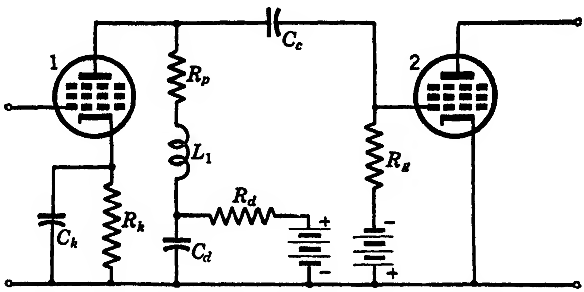


FIG. 19. Circuit for low-frequency compensation.

coupling condenser C_c will usually have several micromicrofarads of capacity to ground and may be placed on whichever side of L_1 gives the better balance.

The other circuits of Fig. 18 permit the use of even greater values for R_p up to the maximum $R_p = 2X_{Ct}$. These circuits are based upon M-derived and constant-K filter networks, and hence have a sharp cutoff characteristic. They are somewhat more difficult to adjust than the series or shunt circuits and are likely to produce unwanted transients in the cutoff region.*

At low frequencies the reactance of the coupling condenser C_c becomes appreciable, and, at the point where this reactance is equal to the grid resistance R_g, gain will be approximately 70 per cent of its mid-frequency value and there will be a 45-degree phase shift. A condenser-resistor combination in series with the plate load, as in Fig. 19, can be made to compensate for the effect of C_c. R_d should be at least ten times the reactance of C_d at the lowest frequency. Compensation results because the signal voltage drop across capacitor C_d varies with frequency directly as it does across C_c. Hence at low

* "Video Output Systems," Foster and Rankin, *RCA Review*, April 1941.

frequencies the plate voltage of tube 1 rises to offset the loss in signal voltage across the capacitor C_c.

If grid bias is obtained from the drop across a cathode resistance, the cathode by-pass condenser should have negligible reactance at low frequencies. When the reduction in amplifier gain can be tolerated, it is often advisable to omit entirely the cathode condenser in the interest of uniform response. Some high-frequency compensation may be obtained through the use of a very small by-pass condenser whose reactance is effective only at the high frequencies.

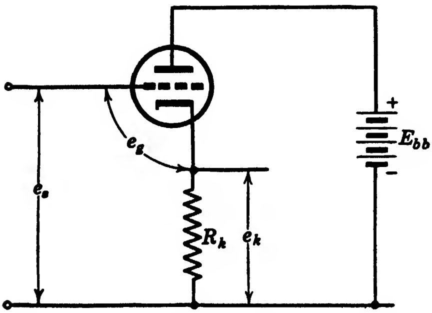


FIG. 20. Cathode-follower circuit.

Cathode Followers. A cathode follower is an amplifier circuit in which the load resistance is placed in the cathode circuit instead of the plate circuit. The name cathode follower suggests that the output signal follows the input signal closely in magnitude, phase, and frequency. A typical circuit for a triode is shown in Fig. 20. The input signal e_s is applied across the grid and the lower side of the cathode load resistor R_k, while the load output voltage e_k appears across the terminals of R_k. Circuit operation may be analyzed by either the constant voltage or constant-current form of equivalent circuit. Under the former

$$e_s = e_g + e_k$$

$$G = \frac{e_k}{e_s} = \frac{e_k}{e_g + e_k}$$

Hence, from equations 5 and 6, Chapter VI,

$$e_k = \frac{\mu e_g R_k}{R_k + r_p}$$

$$G = \frac{e_k}{e_s} = \frac{\dfrac{\mu e_g R_k}{R_k + r_p}}{e_g + \dfrac{\mu e_g R_k}{R_k + r_p}}$$

$$G = \frac{\mu R_k}{R_k + r_p + \mu R_k} = \frac{\mu R_k}{(1 + \mu)R_k + r_p} \tag{17}$$

Also, from equations 8, 9, and 10, Chapter VI,

$$i_p = \frac{g_m e_g r_p}{r_p + R_k}$$

where R_k is the load resistor.

$$e_k = i_p R_k = \frac{g_m e_g r_p R_k}{r_p + R_k}$$

$$G = \frac{e_k}{e_s} = \frac{e_k}{e_k + e_g} = \frac{\dfrac{g_m e_g r_p R_k}{r_p + R_k}}{\dfrac{g_m e_g r_p R_k}{r_p + R_k} + e_g}$$

$$G = \frac{g_m R_k}{\dfrac{R_k}{r_p} + g_m R_k + 1} \tag{18}$$

If $r_p \gg R_k$, then $\dfrac{R_k}{r_p} > 0$, and

$$G = \frac{g_m R_k}{g_m R_k + 1} = \frac{1}{1 + \dfrac{1}{g_m R_k}} \tag{19}$$

It is obvious from either equation 17 or 18 that, for constant μ and g_m, the gain will increase with R_k to the maximum value, one.

While the cathode follower has a gain of less than one, it possesses the following important advantages: (1) It may be used to couple a high-impedance source to a low-impedance circuit without distortion, (2) it provides an output circuit with one side at ground potential, and (3) the output voltage is in phase with the input signal. The reason for these advantages lies in a low input admittance and a simple grounded output circuit. These are illustrated in Fig. 21 which shows that the shunt impedances across the input consist (part *b*) of the grid-to-plate interelectrode capacitance c_{gp} and a fraction (less than one) of the grid-to-cathode capacitance c_{gk}. Here it should be noted that the former capacitance has a factor of unity instead of the $(1 + G)$ value in the conventional amplifier circuit, and the latter capacitance is also less than in the conventional circuit. Thus the input admittance is capacitive but very low in magnitude even for

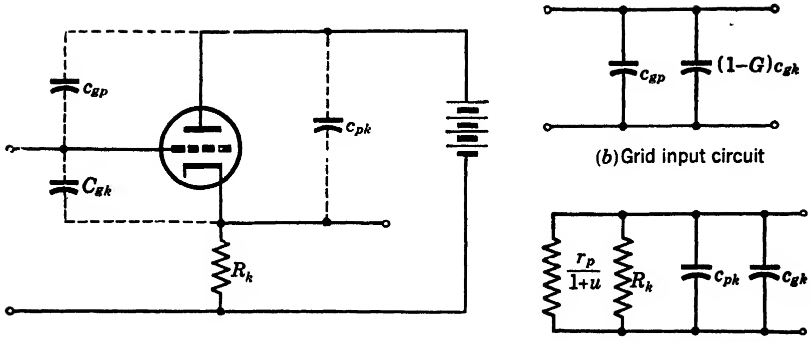


FIG. 21. Equivalent circuits for cathode followers.

high frequencies. The output impedance in Fig. 21c consists of R_k and a fraction of the plate resistance r_p in parallel with the interelectrode capacitances c_{gk} and c_{pk}. This equivalent circuit may be expressed in terms of admittance as

$$Y_0 = \frac{1}{R_k} + \frac{\mu + 1}{r_p} + j\omega(c_{gk} + c_{pk}) \tag{20}$$

Cathode followers perform well at any frequency (even greater than 100 megacycles) where cathode-to-ground capacitance is negligible or may be neutralized.

Grounded-Grid Amplifier. Another method of using a triode as an amplifier is shown in Fig. 22a. This is commonly known as the

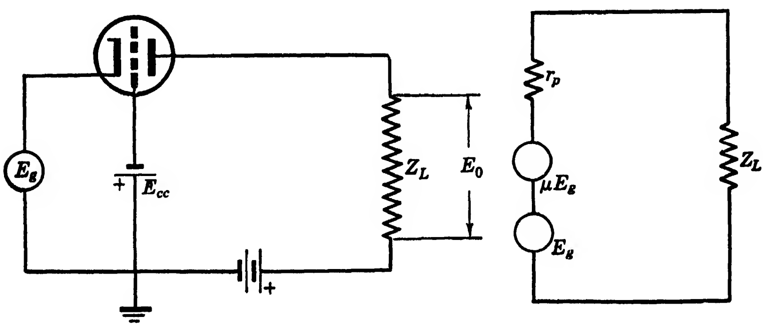


FIG. 22. Actual and equivalent circuit of a grounded-grid amplifier.

grounded-grid amplifier since the grid is at ground potential for alternating current. The input is applied between the cathode and ground, and the output is developed across an impedance Z_L in the plate circuit.

The equivalent circuit for the grounded-grid amplifier is shown in Fig. 22*b*. The internal impedance of the input source is assumed to be zero. The input source E_g and the equivalent generator μE_g are in series. The resulting plate current I_p is given by:

$$I_p = \frac{\mu E_g + E_g}{r_p + Z_L} = \frac{E_g(\mu + 1)}{r_p + Z_L} \tag{21}$$

The output voltage E_0 is given by

$$E_0 = -I_p Z_L = \frac{-E_g Z_L(\mu + 1)}{r_p + Z_L} \tag{22}$$

and the voltage gain is given by

$$\frac{E_0}{E_g} = \frac{-Z_L(\mu + 1)}{r_p + Z_L} \tag{23}$$

The grounded-grid amplifier has slightly more gain than the conventional amplifier provided that the internal impedance of the E_g is very much smaller than $r_p + Z_L$.

The input impedance of the grounded-grid amplifier is given by

$$Z_i = \frac{E_g}{I_p} = \frac{r_p + Z_L}{\mu + 1} \tag{24}$$

This is quite low and requires a driving source E_g of low internal impedance in order to realize maximum gain or develop useful power across Z_L.

The grounded-grid amplifier finds its greatest application in vhf and uhf circuits. The grounded grid acts as an electrostatic shield between the input and output circuits. Thus triodes can be used at vhf and uhf without neutralization. This is very advantageous as satisfactory neutralization is very difficult to obtain at these frequencies. A grounded-grid triode is preferable to pentodes and tetrodes at uhf when the residual inductance in the screen and suppressor leads causes instability. When used as a power amplifier the high driving power is not objectionable as the driving sources act in series with the grounded-grid amplifier to develop across the load Z_L. Optimum performance is obtained with tubes specifically designed for grounded-grid operation, such as the lighthouse tubes (see page 420).

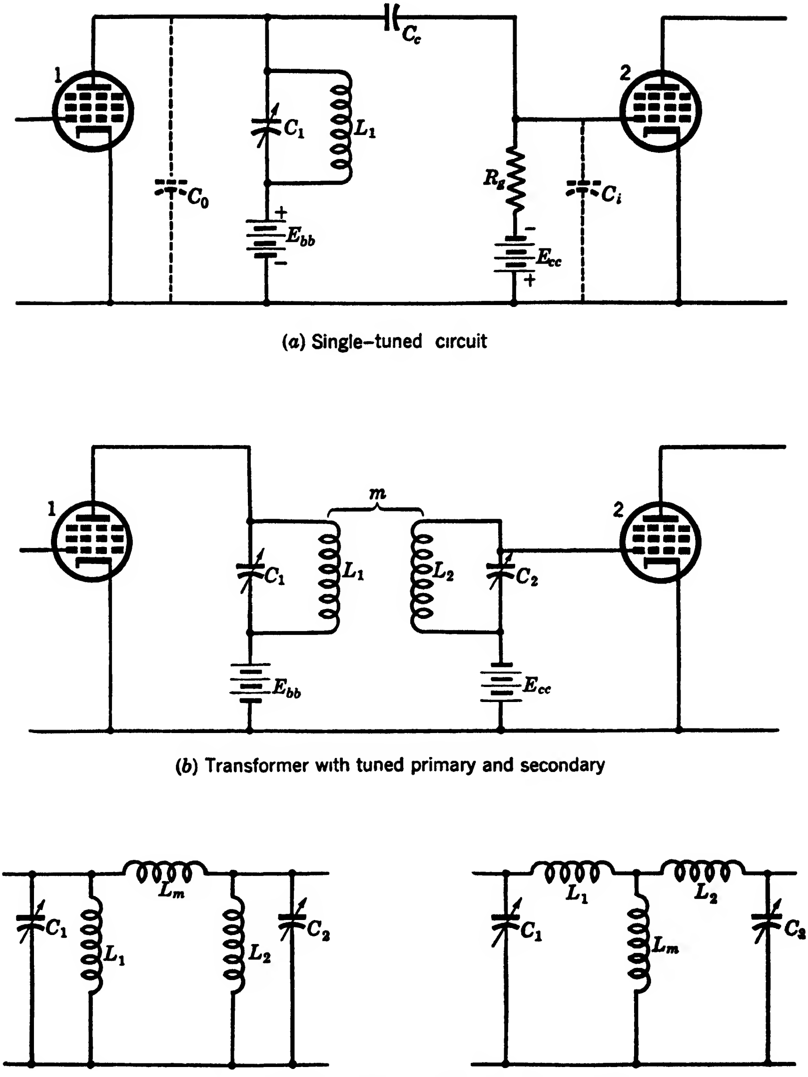
Radio-Frequency Amplifiers. Radio-frequency amplifiers operate above the audible range and are generally designed to respond only

to a narrow "pass band" seldom greater than 20 per cent and usually 1 or 2 per cent of the mid-frequency. Thus they differ from the amplifiers previously discussed which were able to pass a relatively wide band of frequencies. Radio-frequency amplifiers are necessary in radio and carrier telephone transmitters, receivers, and other applications where a narrow band of frequencies must be selected and all others rejected.

There are numerous circuits for coupling rf stages, ranging from the simple tuned circuit of Fig. 23*a* to complicated multisection networks. All circuits employ resonance in some form, and all serve to cancel the shunting effect of the tube and the circuit capacitances. The latter function is essential because at the relatively low frequency of 10 megacycles the shunt reactance of a circuit may become approximately 1000 ohms, a value too low to permit much amplification. In Fig. 23*a* the tuning capacity C_1, output capacity C_0 of tube 1, input capacity C_i of tube 2, and all stray wiring capacities act together with the coil L_1 to form a parallel resonant circuit. Under some conditions C_i varies with the electron current in the tube. This effect, called the Miller effect (see page 136), occurs particularly with high g_m tubes when gain is controlled by varying the grid bias. In order to prevent such change in input capacity and the resulting shift in resonant frequency, a small unby-passed resistance may be inserted in the cathode circuit.

Tubes employing screen grids are more desirable than triodes as radio-frequency or rf amplifiers because they act as constant-current generators and their output voltage therefore varies directly with load impedance. Also, their low grid-to-plate capacity results in low effective input capacity and greater stability. If enough energy is coupled from the plate back to the grid via c_{gp}, the circuit will "oscillate" and will be useless as an amplifier. Such oscillation may be prevented by reducing the gain, increasing the grid circuit losses, or by canceling the feedback with a neutralizing circuit. In Fig. 24*a* the output circuit has a tap at the center. Voltage 180 degrees out of phase with E_p is fed through a small variable condenser c_n back to the grid. c_n is adjusted to be approximately equal to c_{gp} which reduces the net current from plate circuit to grid to zero. The same principle is applied in Fig. 24*b* to two tubes connected in "push-pull." Such a stage is said to be "cross neutralized."

The signal voltage at the plate of an rf stage is equal to $-i_pZ_L$, where Z_L is the impedance into which the tube operates. With screen-



(*a*) Single-tuned circuit

(*b*) Transformer with tuned primary and secondary

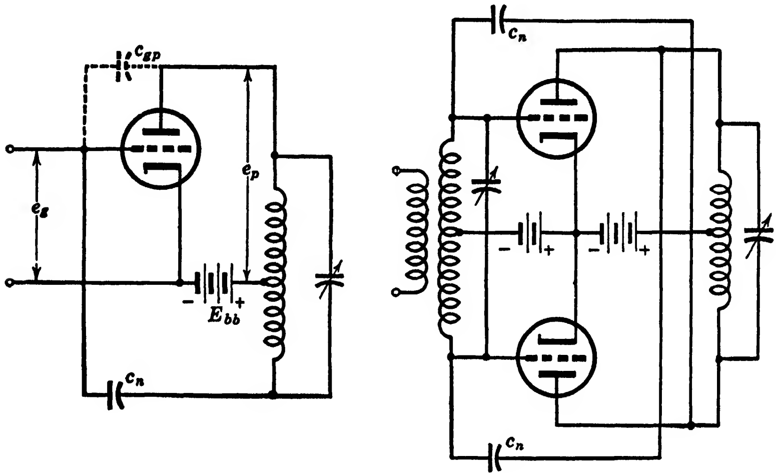
(*c*) Physical mutual coupling in π and T circuits

FIG. 23. Coupling circuits for rf amplifiers.

grid tubes this is approximately $-e_g g_m Z_L$, and, if the voltage at the following grid is the same as that in circuit Fig. 23*a*, stage gain is

$$G = -g_m Z_L \tag{25}$$

Expressions for tuned-circuit impedance were given in Chapter V, and the manner in which this impedance varies with frequency defines the "selectivity" or ability of the stage to discriminate against unwanted



(*a*) (*b*)

Fig. 24. Circuits for neutralization of grid-plate capacitance.

signals. When the effective resistance of the coil is low the tuned-circuit impedance is approximately:

$$Z_L = \frac{L}{R_L C} = QX_c \quad (\textit{See} \text{ equations 9 and 10, Chapter V})$$

Hence, from equation 25,

$$G = -g_m Q X_c \tag{26}$$

Amplifiers that must be tuned to different frequencies, such as those in a radio receiving set, must have either L or C variable. In the more common arrangement where L is fixed and C is a section of a variable condenser gang, the ratio L/C will increase as frequency is squared. Gain will, therefore, be higher at the high-frequency end of the band.

A more economical and in many respects a more satisfactory coupling device is an rf transformer (Fig. 23*b*). Both the coupling con-

denser and grid resistor are omitted, although an additional winding is required. Either primary, secondary, or both may be tuned. An untuned primary can be adjusted so that with the distributed and circuit capacities resonance occurs just below the tuning band. At the low-frequency end of the band where the primary is approaching resonance, gain will be increased. This may be made to compensate for the reduced L/C ratio in the tuned-grid circuit. A high-frequency primary, one which is resonant above the tuned band, is less expensive and affords a better impedance match when the primary is connected to a transmission line. Greater selectivity is obtained through the use of a double-tuned transformer. Such an arrangement is ideal for amplifiers that operate at a fixed frequency, such as the intermediate-frequency amplifiers in a superheterodyne receiver but is unwieldy when the tuning must be varied. Plate voltage can be found as already discussed. Using for Z_L the input impedance of the transformer, e_{g2} will be e_{p1} times the gain of the transformer.

For uniform response over the pass band the coupling may exceed critical so that a flat-topped selectivity curve results. A wider pass band can be provided if loading resistors are connected across primary and secondary coils. This use of resistors reduces the gain so that a greater number of stages will be required. It is sometimes easier to control the coupling in a transformer when a "physical mutual" is used. Then the primary and secondary coils are arranged to have little mutual coupling, and a third coil L_M (Fig. 23*c*) provides the desired voltage transfer. For the π connection L_M must be large compared to L_1 and L_2 and the coupling increases as L_M decreases. In the T circuit the reverse is true with L_M being relatively small and the coupling increasing with L_M.

At the higher frequencies, the tuned circuits may be replaced by quarter-wave transmission line sections. A concentric line or coaxial cable consisting of a small inner conductor supported on insulators at the center of a hollow pipe is well suited to this purpose. If a section of line one-quarter wavelength long has one end shorted, the other end will present a very high impedance. This is the same as the high impedance of a parallel resonant circuit composed of a coil and condenser. A transmission line has the advantage of being simple and mechanically rigid. A three-inch section with a small "trimmer" condenser at the open end will tune to 300 megacycles. Two-line sections may be coupled to form a transformer by connecting a link between the inner conductors at a point near the shorted ends. As this link is moved toward the open ends, the coupling increases.

Problems

1. In the amplifier shown in Fig. 25, $g_m = 1200$ μmhos; $r_p = 1$ megohm; $R_p = 100{,}000$ ohms; $R_g = 1$ megohm; $C_t = 30$ $\mu\mu$f; $C_c = 0.005$ μf. (a) What is the gain for middle-range frequencies? (b) At what frequencies is the gain 0.707 of the middle-range gain?

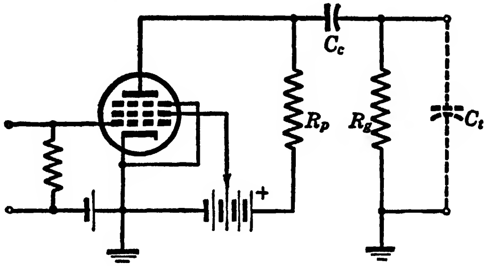


Fig. 25

2. In the amplifier of Fig. 26, $\mu = 90$; $r_p = 83{,}000$ ohms; $R_p = 250{,}000$ ohms; $R_g = 500{,}000$ ohms; $C_c = 0.01$ μf; $C_s = 60$ $\mu\mu$f. (a) Compute the gain for middle-range frequencies. (b) Compute the gain for the following frequencies and plot the response curve as shown in Fig. 7: 10; 20; 30; 60; 100; 10,000; 20,000; 40,000; 60,000.

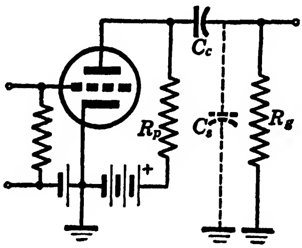


Fig. 26

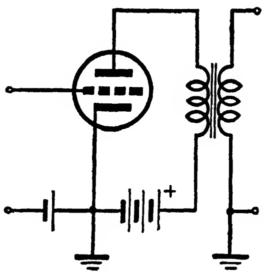


Fig. 27

3. In the amplifier of Fig. 25, $g_m = 1200$; $r_p = 1$ megohm; $R_g = 1$ megohm; $C_t = 12$ $\mu\mu$f. It is desired that the gain shall be down only 0.8 at frequencies of 20 and 150,000 cps. (a) Calculate R_p. (b) What value should C_c have? (c) What is the middle-range gain?

4. A transformer-coupled amplifier is shown in Fig. 27. $\mu = 20$; $r_p = 7000$ ohms; $R_1 = 500$ ohms; $L_1 = 0.1$ henry; $L_m = 40$ henries; $C_{12} = 150$ $\mu\mu$f; $R_2 = 9000$ ohms; $L_2 = 0.9$ henry; $C_2 = 1200$ $\mu\mu$f; $n = 3$. (a) What is the gain for the middle-range frequencies? (b) What is the gain at 10 cps? (c) At what frequency does maximum gain occur? (d) What is the peak gain?

5. In Fig. 28 is shown a two-stage video amplifier. The amplifier is to pass frequencies up to 4 megacycles per second. $C_1 = 8$ $\mu\mu$f (tube capacitance plus

wiring capacitance); $C_2 = 10\ \mu\mu f$ (tube capacitance plus wiring capacitance); $C_3 = 20\ \mu\mu f$ (picture tube capacitance plus wiring capacitance); $c_{gp} = 0.06\ \mu\mu f$ (6AG7); $C_{\text{input}} = 12.5\ \mu\mu f$ (6AG7); $g_m = 2500\ \mu$mhos (6J5); $g_m = 11{,}000\ \mu$mhos (6AG7). Calculate: (a) R_{p2}; (b) L_2; (c) G_2 (gain of 6AG7); (d) R_{p1}; (e) L_1; (f) G_1 (gain of 6J5); (g) over-all gain. (Assume suppressor grid connected to ground.)

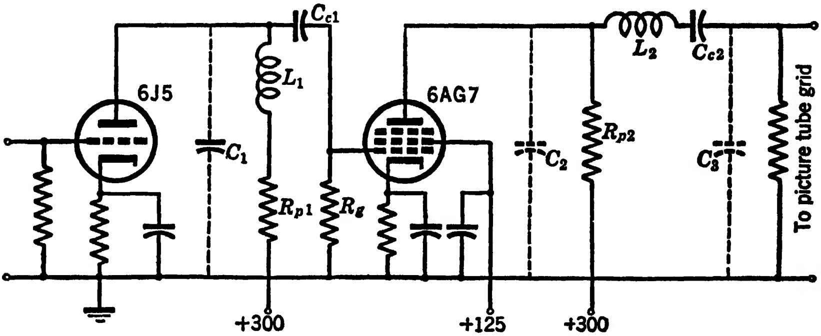


FIG. 28

6. Draw the equivalent circuit of a cathode follower and show that (a)

$$\text{Gain} = \frac{\mu}{\mu + 1} \frac{R_k}{\dfrac{r_p}{\mu + 1} + R_k}$$

(b) Compare this formula with the formula for the gain of an ordinary amplifier. (c) Show that the output resistance of a cathode follower is

$$\frac{R_k r_p}{R_k(\mu + 1) + r_p}$$

7. In the cathode follower of Fig. 29, $\mu = 20$; $r_p = 8000$ ohms. (a) What is the gain? (b) What is the output resistance?

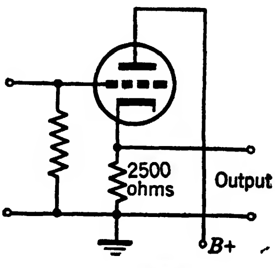


FIG. 29

8. A 6J4 is to be used as a grounded-grid rf amplifier. The plate-load circuit consists of a tuned circuit whose impedance is 50,000 ohms. $\mu = 55$; $r_p = 4500$. Calculate: (a) the voltage gain; (b) the input impedance.

9. A 6SJ7 tube is used in a tuned rf amplifier as shown in Fig. 30. $E_{bb} =$ 100 volts; $L = 0.18$ mh; $f = 1000$ kc; $Q = 175$; $r_p = 0.7$ megohm; $g_m = 1575$ μmhos. (a) Draw the equivalent circuit and calculate the voltage gain. (b) Calculate the gain using equation 11 and compare with your answer to part (*a*).

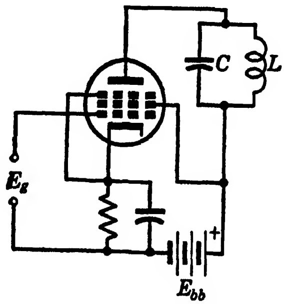


FIG. 30

10. Calculate the gain of a three-stage video intermediate-frequency (if) amplifier using 6AC7 tubes. The if transformers are loaded on the secondary side with resistors so that the plate load for each stage is approximately 2000 ohms over the pass band. The voltage gain of the transformers is unity.

REFERENCES

M.I.T. ELECTRICAL ENGINEERING STAFF, *Applied Electronics*, John Wiley & Sons, 1943, pp. 490–504.

TERMAN, *Fundamentals of Radio*, McGraw-Hill, 1938, pp. 104–113.

Chapter VIII

POWER AMPLIFIERS

The amplifiers discussed in the preceding chapters are designed to increase the voltage of a weak signal and hence are called "voltage amplifiers." The power amplifier is designed to deliver energy to a load such as a loudspeaker, a transmitting antenna, or diathermy patient, and similar applications. Here current is important as well as voltage, and attention must be given to additional factors such as maximum cathode emission and safe plate dissipation. Power amplifiers are classified as A, AB, B, and C, according to their grid-voltage plate-current excursions, and these different classifications will be discussed in order.

Class A amplifiers are those in which the grid bias and a-c plate voltages are such that plate current never reaches zero at any portion of their operating cycle.

The operation for Class A is shown in Fig. 1. The grid is operated in the negative region between cutoff and zero. In general, for single-tube use the grid bias is selected so that operation takes place on the linear portion of the dynamic transfer characteristic. Here the distortion generated by the tube is low, and low distortion is the outstanding advantage of Class A amplifiers. The requirements for low distortion produce a low plate efficiency. The plate efficiency of an amplifier is the ratio of the a-c (fundamental) plate output to the plate input where the plate input is the product of the plate-supply voltage and the average plate current. The plate efficiencies for Class A operation are of the order of 20 to 25 per cent. With no grid signal all the power delivered by the B supply goes toward heating in the plate circuit of the tube. When signal is applied, some of this power is converted into useful output and the heating of the tube is less. In other words, the tube of a Class A amplifier will be cooler when delivering power to the load than when it is idling. A tube should be chosen which has a rated plate dissipation equal to about three times the desired output power. Peak cathode emission requirements are not severe since the plate current never reaches more than twice the average value.

Maximum power output from any generator is obtained when the load resistance is equal to the internal resistance of the generator. This is true also of the Class A amplifier. However, it is necessary to operate a triode with a load at least twice the plate resistance in order to keep distortion to a reasonable level. The major distortion component for the triode is the second harmonic which can be found from the increase in average plate current with signal as explained in Chapter V. Harmonic distortion is usually expressed as a percentage

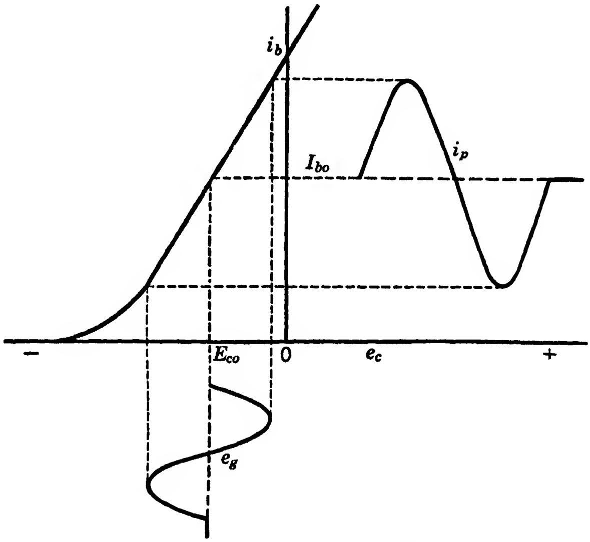


FIG. 1. Class A amplification.

of the total current. In Fig. 16 of Chapter V, the maximum of the current variations is $\frac{1}{2}(I_{max} - I_{min})$. Then using the expression for the second harmonic given in equation 19 of Chapter V, the per cent of harmonic distortion may be resolved into the following equation:

$$\% \text{ 2nd harmonic} = \frac{\frac{1}{2}(I_{b\ max} + I_{b\ min}) - I_{b0}}{(I_{b\ max} - I_{b\ min})} \times 100 \qquad (1)$$

The i_b-e_b characteristics for pentodes and beam-power tubes curve oppositely from those of triodes, and optimum operation is obtained when the plate load is relatively low (approximately one-tenth the plate resistance). The load line should not cross $E_{co} = 0$ below the knee of the curve or distortion and screen current will rise rapidly and the safe screen dissipation may be exceeded. Second harmonics are usually low but the third and fifth become appreciable. Distor-

tion in pentodes may be calculated approximately by the "5-ordinate" method.*

$$\% \text{ 2nd harmonic} = \frac{I_{b\,\max} + I_{b\,\min} - 2I_{b0}}{I_{b\,\max} - I_{b\,\min} + 1.41(I_x - I_y)} \times 100 \qquad (2)$$

$$\% \text{ 3rd harmonic} = \frac{I_{b\,\max} - I_{b\,\min} - 1.41(I_x - I_y)}{I_{b\,\max} - I_{b\,\min} + 1.41(I_x - I_y)} \times 100 \qquad (3)$$

where I_x and I_y are the plate currents corresponding to 0.293 and $1.707E_o$, respectively.

Power output and optimum load resistance can be most easily determined graphically. For triodes and pentodes having small third-harmonic distortion, equation 13, Chapter VI, may be used. When the third harmonic is appreciable, greater accuracy is given by

$$\text{Average power} = \frac{[I_{b\,\max} - I_{b\,\min} + 1.41(I_x - I_y)]^2 R_L}{32} \qquad (4)$$

where I_x and I_y are as defined above.

Power stages are nearly always transformer-coupled since this coupling provides low loss for the steady component of plate current and permits matching the tube to a load of any impedance. Transformer losses usually run between 7 and 20 per cent and must not be overlooked when calculating output. All the power equations give tube output and not useful power in the load.

Class A stages may be used for audio-, video-, or radio-frequency amplification. The power stage of a radio receiver having a single output tube is always Class A, and the preceding radio-frequency and audio-voltage amplifiers are also Class A.

If more power is required than can be obtained from a single tube, two or more similar tubes may be operated in parallel. Load resistance, output, and distortion are found the same as for a single tube except that the current values on the i_b-e_b curves must be multiplied by the number of tubes that are in parallel. Circuits of this type are likely to generate very high-frequency "parasitic" oscillations which must be suppressed by the use of small resistors at the grid of each tube.

A preferable method for utilizing two tubes having identical characteristics is the push-pull circuit shown in Fig. 2. The grids are

* *Radiotron Designer's Handbook,* 3rd ed., pp. 283–284, The Wireless Press, Sydney, Australia.

driven by voltages equal in amplitude but opposite in polarity so that the plate current in one tube increases while that in the other decreases.

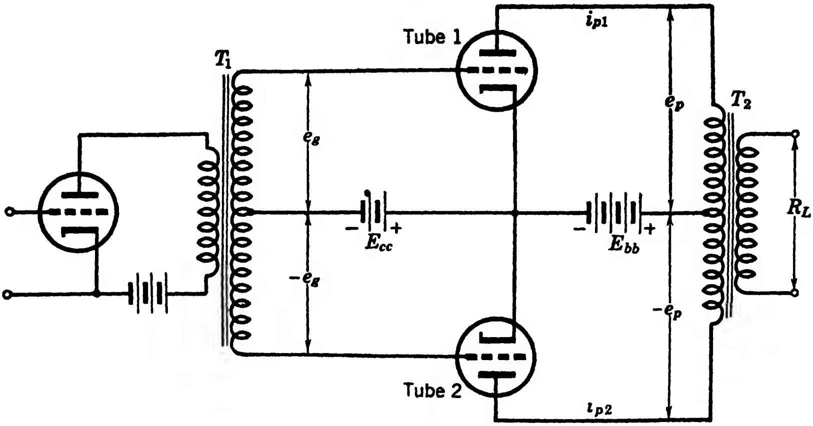


FIG. 2. Push-pull circuit for amplification.

The result of this action is to greatly reduce the second and all other even harmonics so that for a given power output much less total distortion is introduced. The steady component of plate electron current

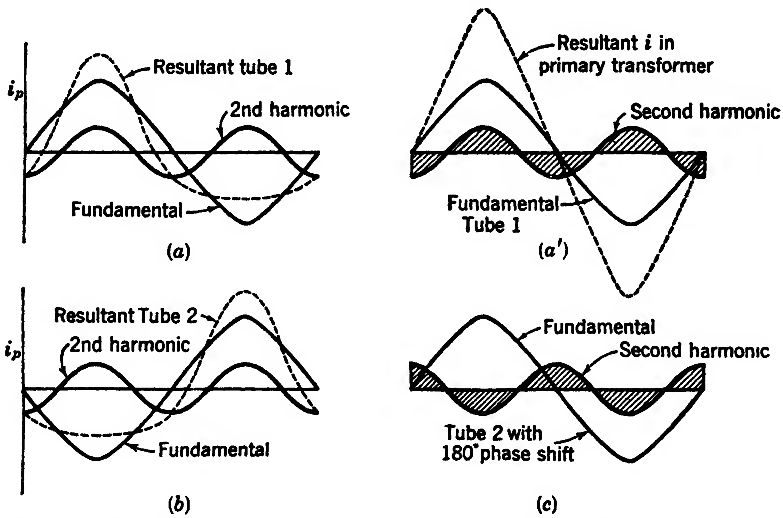


FIG. 3. Reduction in harmonic distortion in push-pull Class A operation.

flows from each end of the winding toward the center tap. If the two halves of the winding are identical and the two plate currents are

equal, the total magnetizing force will be zero. Thus an appreciable saving in core size is made, and the ability of the transformer to pass low frequencies is retained.

The reason for the low harmonic distortion in the push-pull amplifier circuit is illustrated in the curves of Fig. 3. Since operation for wide swings of grid voltage will take place over nonlinear sections of the dynamic plate curves, nonsymmetrical current waves will exist in the individual tube circuits as shown by the resultant waves of parts *a* and *b*. The second harmonic (peak) will be in phase with the peak of the positive loop as shown, and hence in phase with each other when considered within the individual tube circuits. The push-pull connection to the output transformer reverses (180-degree phase shift) the phase relation of tube 2 with respect to tube 1. Thus curves of *a'* are a replica of *a*, whereas curves of *c* are reversed in phase. Inspection of parts *a'* and *c* of Fig. 3 shows that the second harmonics from tubes 1 and 2 flowing through the transformer primary are 180 degrees out of phase, and hence cancel, whereas the fundamental current waves from the tubes are in phase and therefore add, giving the resultant (2 times fundamental) shown by the dotted line in *a'*. Slight differences in tube characteristics and circuit adjustment will prevent a perfect cancellation of second harmonic though the distortion will be low.

The advantages of Class A push-pull amplification may be summarized as follows:

(*a*) No d-c saturation of the output transformer core; hence reduced size and cost.
(*b*) Reduced even harmonic distortion.
(*c*) Insensitive to ripple voltages present in the plate, grid, and filament supplies.
(*d*) No reaction of signal current on the plate power supply; hence no coupling with other stages through common supply leads.
(*e*) Greater power output because of reduced effective plate resistance.

Disadvantages are:

(*a*) Balanced input signal voltage required.
(*b*) Increased cost due to use of two tubes, two sockets, and center-tapped transformer. (Partly offset by saving in transformer core size.)

The reason for the advantages listed as *c* and *d* in the preceding arises from the fact that the a-c components of plate current for the two tubes are equal in magnitude and 180 degrees out of phase. Thus, $i_{p1} = -i_{p2}$ in Fig. 2. Accordingly, the effects of ripples in the supply voltages for plate, grid, and cathode will cancel out in the primary of

the output transformer. The preceding relationship indicates that the current delivered by E_{bb} is constant and independent of signal changes.

The reason for advantage e in the preceding may be learned through a deductive process. Assume first that the turns ratio for the output transformer in Fig. 2 is 2/1 and the secondary load is R_L. Since the turns ratio for each half primary winding to secondary is 1/1, *the load for each tube looking into the transformer is R_L.* Since the total primary turns from plate to plate is two times that of one tube and since the impedance varies as the square of the turns, *the plate-to-plate load is $4R_L$.* The output a-c current for tube 1 (Fig. 2) is i_{p1}, and it flows through one-half of the primary winding. The same ampere-turns may be produced by one-half i_{p1} flowing through the entire primary winding. Continuing this concept, the equivalent current through the entire primary is $\frac{1}{2}(i_{p1} + i_{p2})$ which is equal to i_{p1}, since i_{p1} and i_{p2} are equal in magnitude, 180 degrees out of phase, and fed to the primary in opposite directions. Since the plate-to-plate load resistance is $4R_L$, the a-c voltage drop across the primary is $4i_{p1}R_L$, and that across the load for each tube is $2i_pR_L$. This value is 2 times the voltage drop for a single tube operation where current is i_{p1} and load R_L. This value can be produced by a current $2i_{p1}$ flowing through a load resistance R_L. Thus it may be deduced that for the case assumed the equivalent plate resistance has been reduced to one-half that for single-tube operation, a larger current being thus permitted to flow.

Operating conditions for push-pull Class A amplifiers can be determined graphically as with a single tube. Here, however, a composite i_b-e_b characteristic for the two tubes should be constructed.* This is done by inverting the family of curves for one tube and placing it adjacent to the other so that the operating plate voltages coincide for the chosen value of E_{bb} as illustrated in Fig. 4. Composite operating curves are now constructed for corresponding grid voltages which fulfill the requirements stated on the curve sheet. Requirement number one is met if the curves have been inverted properly. Thus point 100 (e_b) on the zero-zero axis for the upper curves corresponds to 500 for the lower curves. The sum of these values $100 + 500 = 600$ which checks the value of $2 \times 300 = 600$ for the supply voltage. Under requirement number two, $2 \times E_{cc} = -120$ volts. Now, if it is desired to construct the composite curve for $E_{cc} = -20$ (upper curves), the corresponding grid curve for the lower group is found by substitution

*B. J. Thompson, "Graphical Determination of Performance of Push-Pull Audio Amplifiers," *Proceedings I.R.E.*, Vol. 21, p. 595, April 1933.

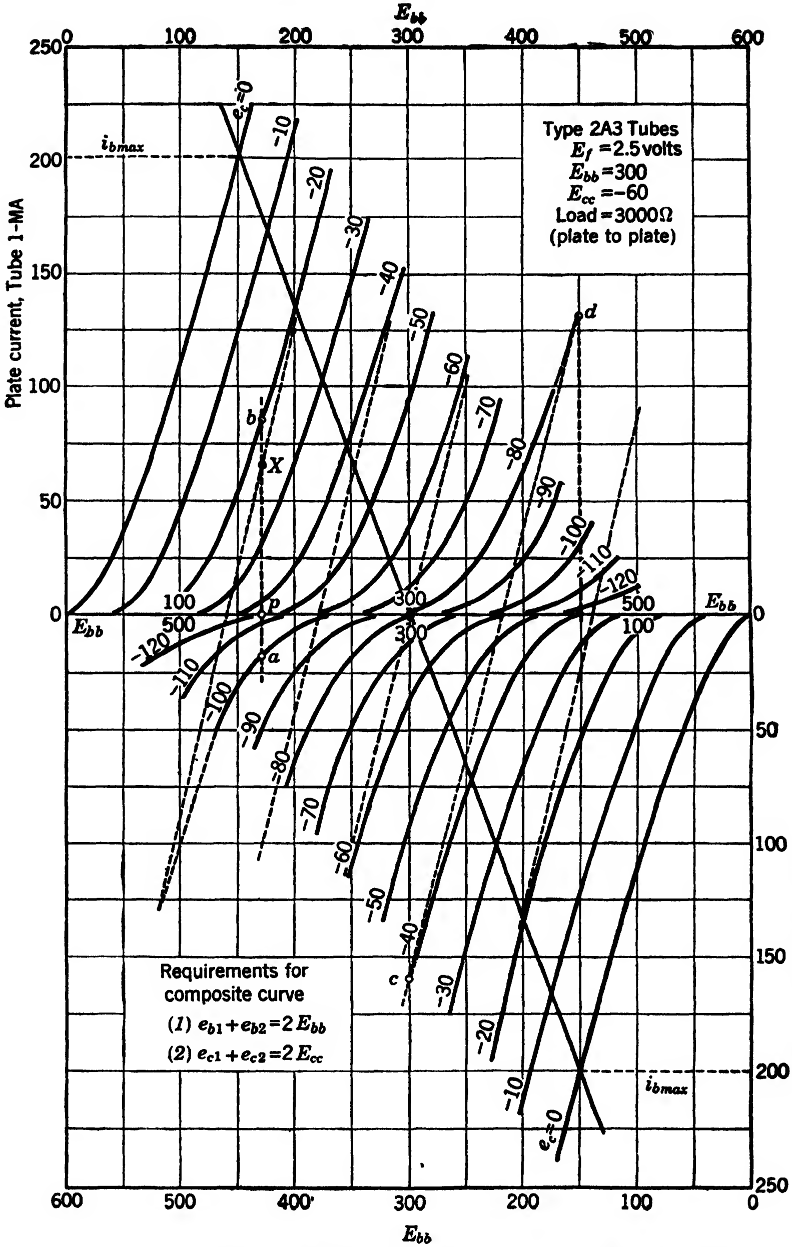


FIG. 4. Composite characteristics for push-pull Class A amplifier.

in the second equation where $-20 + e_{c2} = -120$ and $e_{c2} = -100$ volts. Thus the -20 and -100 are corresponding static curves. The composite curve may now be constructed by erecting a perpendicular line *ab* at any convenient point such as *p*. The intercept *pb* is the plate current i_b of one tube and *pa* the current of the other. Since the latter current is negative, the difference *pX* is the effective current and point *X* is one point on the composite current. Similar graphical constructions at other positions will give additional points for determining the composite curve. An approximate method for determination of the composite curve (for Class A operation *only*) consists in erecting perpendicular lines at the cutoff points of the corresponding grid static curves and then joining the points of intersection of the perpendicular with the other grid curve (see line *cd* as an example).

The load line is drawn so that it crosses the $i_b = 0$ axis at E_{bb}. This line has a slope of $-(1/R_L)$ where R_L is the load looking into one-half the primary of the output transformer. The average power output of a push-pull amplifier in the absence of odd harmonics is

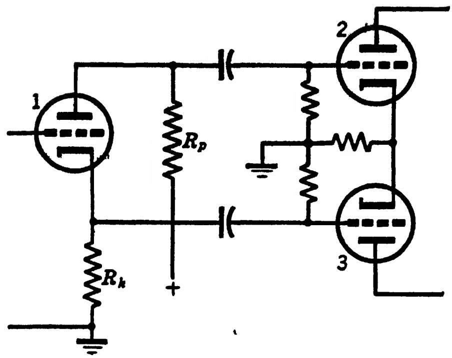
$$\text{Average power} = \frac{(i_{b\,\max})^2 R_L}{2} \tag{5}$$

The actual values of R_L and E_{cc} may vary considerably and should be selected to give maximum power output while the tubes are permitted to operate within the linear region of the composite characteristic and within their rated plate dissipation. A convenient procedure for Class A triodes which must deliver maximum power is to make the load line cross $E_{cc} = 0$ at $0.6E_{bb}$.* E_{cc} should not have a value more negative than one-half that required to give plate-current cutoff at $1.4E_{b0}$ if strictly Class A operation is to be retained.

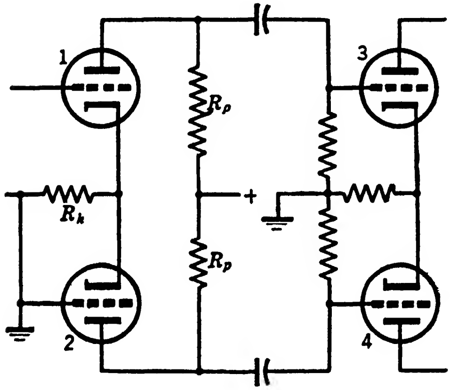
The balanced input voltages required by push-pull stages may be obtained from a grid transformer T_1, Fig. 2, having a center-tapped secondary or from *RC*-coupled "phase inverters," Fig. 5. The plate-cathode loaded inverter *a* has the advantage of simplicity and low distortion. It requires only one tube but has a gain of less than unity. For perfect balance, $R_k = R_p$. Circuit *b* is self-balancing like *a* and, although requiring two tubes, does provide some gain. R_k should be large for best balance, preferably approaching R_p.† In *c*, grid 2 is

* *Receiving Tube Manual RC-14*, Radio Corporation of America, Harrison, New Jersey.

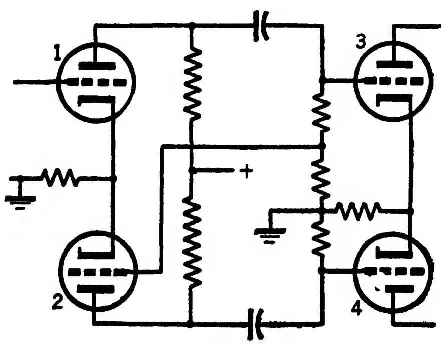
† Wheeler, "Self-Balancing Phase Inverters," p. 67, *Proceedings I.R.E.*, Vol. 34, February 1946.



(a) Plate-cathode coupling



(b) Common cathode coupling



(c) Cascade

FIG. 5. Phase inverters coupled to push-pull stage.

fed a portion of the voltage output of tube 1. The divider must be rather carefully adjusted and, if the tubes are identical, should have a ratio of $1/G_2$ where G_2 is the gain of tube 2. Operating parameters must be properly chosen since any frequency or harmonic distortion created by tube 1 will be amplified by tube 2. Since the grids of Class A output stages are always negative, they draw no current and the driving stage needs to supply only the required signal voltage.

A Class AB amplifier operates with greater negative bias than the Class A as shown in Fig. 6 and thus permits higher plate voltage

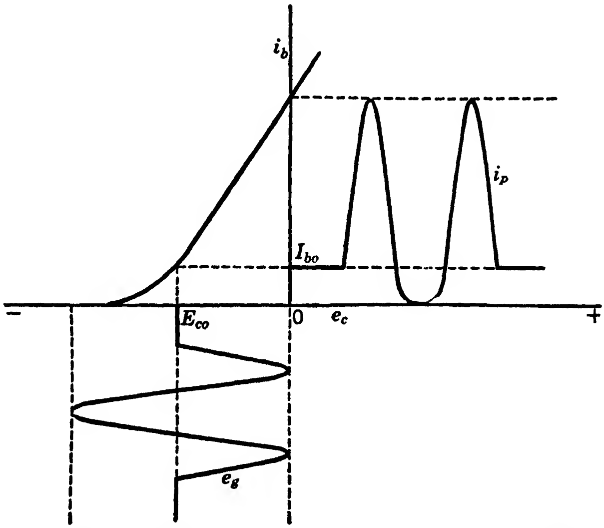


FIG. 6. Grid voltage-plate current operation under Class AB_1.

without exceeding the rated dissipation and greater grid swing; hence power output is increased. Average plate current increases with signal, and the i_p wave is no longer the same form as e_g. Two tubes must be used in push-pull for audio amplifiers or wherever fidelity of output wave form is necessary. This class of operation falls between A and B. It is usually subdivided into AB_1 amplifiers which draw no grid current, and AB_2 in which the grids do swing positive.

The reduction in distortion by push-pull operation for AB amplification is illustrated in Fig. 7. Since the grid signal swings the grid below the cutoff point, the plate current goes to zero for part of each cycle for each tube. The resulting plate currents for the two tubes will be approximately as shown in Fig. 7. The plate currents represented by the areas *aob* and *cod* are equal and opposite. When they

pass through the primary of the transformer, they should neutralize each other and leave a net resultant shown by the dotted line *ac* which means low distortion. The efficiency for AB operation should be higher than for A operation because the a-c output current is of larger magnitude and because the average d-c current from the plate supply is lower in magnitude under similar conditions of operation.

Two typical beam-power tubes (type 6L6) will deliver 17.5 watts when used for Class A operation, 26.5 watts for Class AB_1, and 47 watts with Class AB_2. For AB_2 operation the driver must supply 0.27 watt for the grids with good regulation. To prevent change in

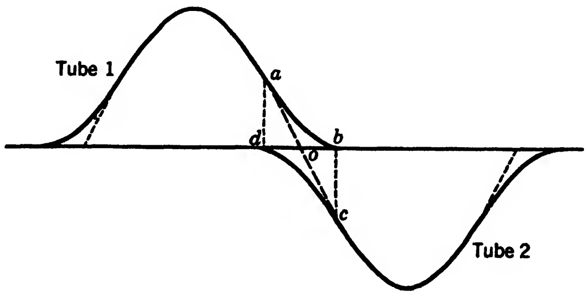


FIG. 7. Cancellation of distortion in push-pull circuit for Class AB amplification.

operating bias arising from grid current flow, the d-c resistance of the grid circuits must be kept at a low level by using transformer or impedance couplings. Since the average plate current varies with signal the plate power supply must have good regulation to maintain constant voltage with and without signal.

Class B amplifiers are those operating with approximately cutoff bias and with zero or very low quiescent plate current as shown in Fig. 8.* Current flows only during the positive half of the e_g cycle, and so for audio work two tubes must be used in push-pull—one to supply each half of the wave. With no signal little power is drawn from the plate supply, a particularly desirable feature for battery operation or for very high-power amplifiers such as those in radio transmitters. The theoretical maximum plate efficiency for this circuit is 78.5 per cent.

This theoretical efficiency may be determined by reference to Fig. 9. The average value for the direct current for a full cycle is $2I_m/\pi$ (cur-

* L. E. Barton, "High Audio Power from Relatively Small Tubes," *Proceedings I.R.E.*, Vol. 19, July 1931.

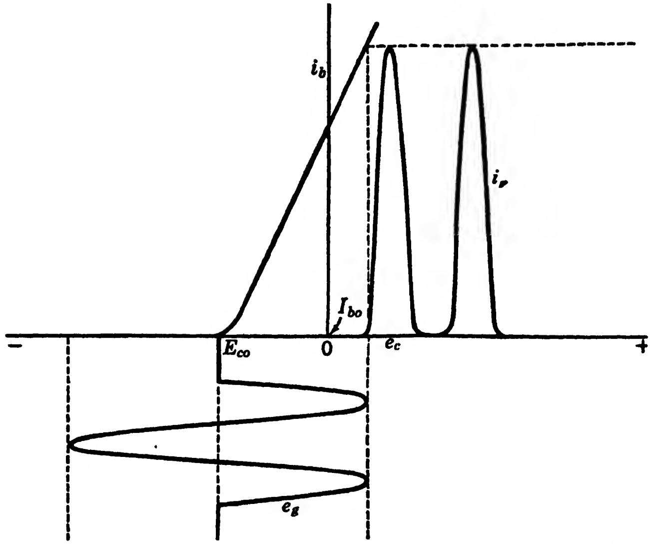


FIG. 8. Grid voltage-plate current Class B amplification with single tube.

rent from two tubes passes plate supply). Thus the plate input from a battery is:

$$P_{\text{input}} = E_{bb} \times \frac{2}{\pi} I_{\max}$$

The effective plate-current output (fundamental) is $I_{\max}/\sqrt{2}$. The maximum voltage swing, as shown by the load line on left of Fig. 9, is $(E_{bb} - E_{\min})$ and the extreme value that it could have, when $E_{\min} = 0$, is E_{bb}. The effective value of this extreme is $E_{bb}/\sqrt{2}$. Thus the theoretical limiting power output is

$$P_{\text{output limit}} = \frac{E_{bb}}{\sqrt{2}} \times \frac{I_{\max}}{\sqrt{2}} = \frac{E_{bb}I_{\max}}{2}$$

from which the theoretical efficiency is

$$\text{Efficiency}_{\max} = \frac{\dfrac{E_{bb}I_{\max}}{2}}{E_{bb} \times \dfrac{2I_{\max}}{\pi}} = \frac{\pi}{4} \quad \text{or} \quad 78.5\% \tag{6}$$

The usual efficiency for Class B audio amplifiers lies within the range of 50 to 60 per cent per stage. This represents a great improve-

ment over Class A efficiency and, in addition, a greater power output can be realized from the same size of tubes. Somewhat more distortion is generated, and this has its greatest magnitude at low signal levels because nonlinearity is more pronounced near cutoff. The restrictions on plate-supply regulation, grid-circuit resistance, and driver impedance hold in even greater degree than for Class AB_2 amplifiers. A Class A power stage coupled through a step-down transformer is generally used for the driver because of its minimum distortion characteristic. Cathode resistance bias is not practical because of the

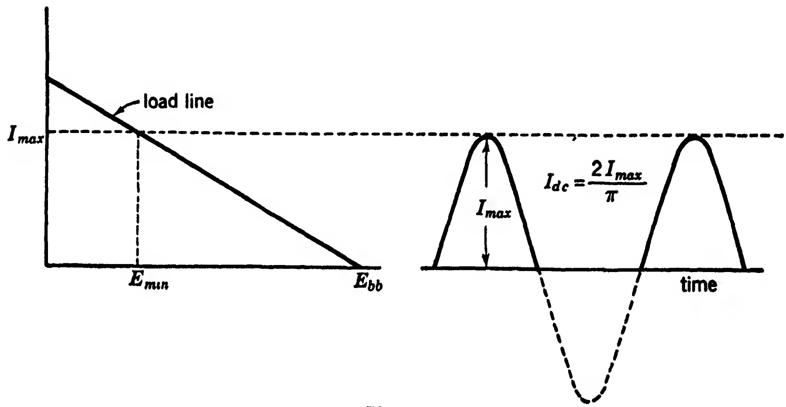


FIG. 9

great current fluctuations and because the bias supply must have low impedance. Tubes particularly well suited to Class B operation are high-μ triodes whose plate current with $e_c = 0$ is low. Such tubes are operated with zero bias supply and present a more nearly constant load for the driver stage.

Design of Class B stages is quite involved because of the many variables. For most purposes data published for the particular tubes being used will give satisfactory values of load resistance, plate voltage, driving power, etc. Raising the load resistance makes the stage easier to drive and increases the efficiency but lowers the power output. The load should be adjusted so that the minimum plate voltage is never less than the maximum positive grid voltage; otherwise the safe grid dissipation may be exceeded. If $e_{b\,min}$ of the load is known or can be estimated, then the other circuit parameters can be found easily by assuming plate current loops of half-sine waves.* Each tube draws current for one-half cycle and during the other half is an infinite im-

* Terman, *Fundamentals of Radio,* p. 168, McGraw-Hill.

pedance.* This is the same as a single tube working for a full cycle into R_L, the impedance of one-half the transformer primary.

$$e_{b\,\min} = E_b - I_{p\,\max}R_L \tag{7}$$

The plate-to-plate impedance is $4R_L$ as in Class A operation.

$$\text{Average power} = \frac{(I_{p\,\max})^2 R_L}{2} \tag{8}$$

Class B stages are often used for rf amplification and may be either push-pull or single-ended. Tuned transformers are always used at radio frequency so that the load impedance to all frequencies other than the fundamental is very low and harmonics are suppressed. A Class B amplifier may be adjusted so that the rf output voltage varies directly with input voltage. Such a stage is called a *linear amplifier* and is useful where a modulated carrier must be amplified. Peak efficiency is the same as for Class B audio but at the unmodulated carrier level the efficiency is only 25 to 33 per cent.

Class C amplifiers have a negative grid bias greater than cutoff and plate current flows for less than one-half cycle as illustrated in Fig. 10. The grid is driven far into the positive region so that it draws

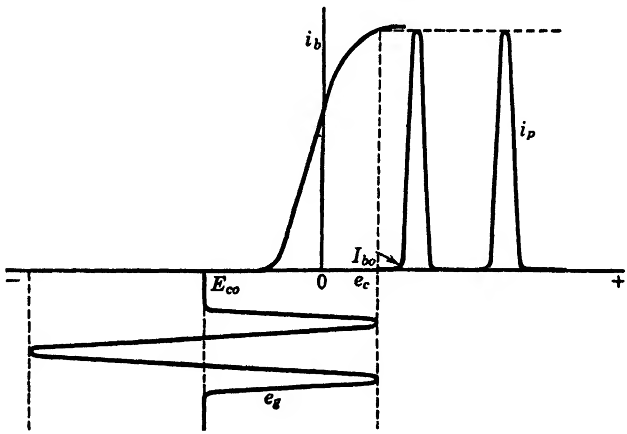


Fig. 10. Grid voltage-plate current amplification with Class C.

considerable current. Usually the bias is 2 to 4 times the cutoff value, and the plate current flows for about one-third of a cycle. Consequently, the current wave is greatly distorted and may reach a high peak value, approximately that of the total cathode emission. Such

* Dow, *Fundamentals of Engineering Electronics*, p. 309, John Wiley & Sons.

an amplifier should be used only where the harmonics are suppressed by tuned circuits. Efficiency is higher than for any other class of amplifier and may exceed 75 per cent in practical applications. This high efficiency makes Class C operation desirable wherever high power is required, such as in radio transmitters, diathermy, and industrial high-frequency heating equipment.

A properly adjusted Class C stage acts like a pure resistance load on the plate supply. If E_{bb} is increased, I_b (d-c component of plate current) will increase in a like ratio and power output will be proportional to $(E_{bb})^2$. Thus a convenient method for modulating the amplitude of the rf carrier is provided (see Chapter X).

Several graphical and algebraic methods for analysis of Class C amplifiers have been proposed.* One of these may be used if the amount of work involved can be justified. It is often sufficiently accurate to assume a plate efficiency ranging from 50 per cent for small tubes to 75 per cent for larger ones. Values for maximum plate voltage, plate dissipation, and grid-driving power are supplied by the tube manufacturer. For example, a type 833A high-frequency triode has the following recommended maximum operating values: unmodulated d-c plate voltage, 4000; d-c plate current, 500 ma; and plate dissipation, 400 watts. If the efficiency is assumed to be 75 per cent, then 25 per cent of the input must be lost on the plate. Permissible d-c input is:

$$P_i = \frac{400}{0.25} = 1600 \text{ watts}$$

Also

$$P_i = I_b E_b$$

$$I_b = \tfrac{1600}{4000} = 0.4 \text{ ampere}$$

Or, if I_b is maximum instead of E_b,

$$E_b = \frac{1600}{0.5} = 3200 \text{ volts}$$

Hence the tube may be operated anywhere between 4000 volts, 400 ma, and 3200 volts, 500 ma, with an input of 1600 watts and an output of 1600 − 400 = 1200 watts. Not all of this output is available for there

* *Applied Electronics,* M.I.T. Electrical Engineering Staff, p. 571, John Wiley & Sons; Everitt, *Communication Engineering,* p. 565, McGraw-Hill; F. E. Terman and W. C. Roake, "Calculation and Design of Class C Amplifiers," *Proceedings I.R.E.,* Vol. 24, p. 620, April 1936.

are some losses in the tuned plate circuit. These are usually small, however, compared to those in the tube.

The grid voltage for zero plate current can be found from the i_b-e_b curves or for triodes may be taken as approximately $-(E_b/\mu)$. Since μ for the type 833 is 35, E_c for Class C operation should be

$$-\frac{2E_b}{\mu} = -\frac{2 \times 4000}{35} = -228 \text{ volts}$$

Grid-driving power (from the manufacturer's data) is approximately 26 watts, although the driver should be able to deliver more than this to supply grid circuit losses and provide a reasonable additional allowance. Insufficient grid drive results in lowered efficiency and loss of the linear relationship between E_b and I_b.

Either triodes or screen-grid tubes, such as tetrodes and pentodes, are suitable as Class C amplifiers. The latter require much less grid-driving power and consequently show greater *power gain* per stage. Also their plate-to-grid capacity is usually low enough so that no neutralization is necessary to prevent oscillation. Careful circuit adjustment must be maintained, however, so that the screen does not overheat.

In passing, it should be noted that the push-pull circuit may be used for Class C power amplification. Distortion will be much greater than with Class A or Class B. The usual application for Class C is with an rf narrow band circuit where the tuned circuit or tank stores energy in its electric and magnetic fields so as to supply both halves of the output waves, as explained in Chapter V. Class B may also be used without the push-pull circuit for radio-frequency amplification where the tuned load impedance supplies the second half-wave for each cycle.

Nothing has been said concerning the value of load impedance for rf amplifiers although the load governs to a large extent efficiency, power output, and harmonic distortion. Discussion of the design of tuned load circuits, including operating Q and L/C ratio, is beyond the scope of this book but may be found in the references already cited and in *Technical Manual TT3*, "Air-Cooled Transmitting Tubes," published by the RCA Manufacturing Company.

Multistage Amplifier Design. The design of a complete amplifier should be approached in a logical manner in order that all factors may be properly evaluated. Specific circuits will be suggested by the general requirements. For example, in designing an audio amplifier it is necessary to consider:

(*a*) Power output required and the type of load which the amplifier must feed, such as a loudspeaker, recorder, Class C rf stage, etc.

(*b*) Voltage gain, which will be governed by the signal source such as microphone, photocell, or others. It may be necessary to provide two or more input circuits so that several signals may be combined.

(*c*) Frequency range, which should be kept to the minimum consistent with the application. Greater range increases the hum, noise, and distortion, whereas, if these are kept low, the cost of the amplifier is increased.

(*d*) Special requirements, such as battery operation for portable use, small size, and others. Limits may also be placed on the number and type of tubes which may be used or on the over-all cost.

After the above factors have been determined, the output stage may be designed. A necessary aid is a chart or handbook showing the characteristics of all available tubes.* Tubes should be selected that will provide the required output with a reasonable safety factor. Here it is important to take into account the losses in the output transformer which usually run from 7 to 20 per cent. If minimum distortion is required or if the output must be greater than can be supplied by a single tube, a push-pull stage should be used. Triodes have the advantage of low plate resistance; hence their output voltage is more nearly constant with variation in load impedance. The type of output tubes and the required power will determine the operating load line, and this factor together with the load resistance dictates the turns ratio for the output transformer. Where large amounts of power are required or where it is necessary to keep the $B+$ supply drain to the minimum, a Class B stage may be used.

After the operating conditions for the output stage are fixed the preceding or driver stage may be designed. For push-pull output this stage may be transformer-coupled or, with negative grid tubes, may be a phase inverter. The driver must provide sufficient voltage swing and, for Class B output stages, sufficient power to drive the final grids for full output. If the signal voltage necessary on the driver grid is less than that supplied by the signal source, additional voltage amplifier stages must be used. These stages may include volume controls, mixing circuits, and tone controls and may use either triodes or screen-grid tubes. Triodes are simpler to connect since they require no screen

* *Tube Handbook HB-3* or *Receiving Tube Manual RC-14*, Radio Corporation of America, Harrison, N. J.; *Technical Manual*, Sylvania Electric Products, Inc., Emporium, Pa.

supply but usually provide lower gain and a poorer high-frequency response.

Amplifiers capable of passing low audio frequencies are subject to an instability known as "motor boating" unless suitable precautions are taken. Motor boating occurs when a plate or screen voltage supply

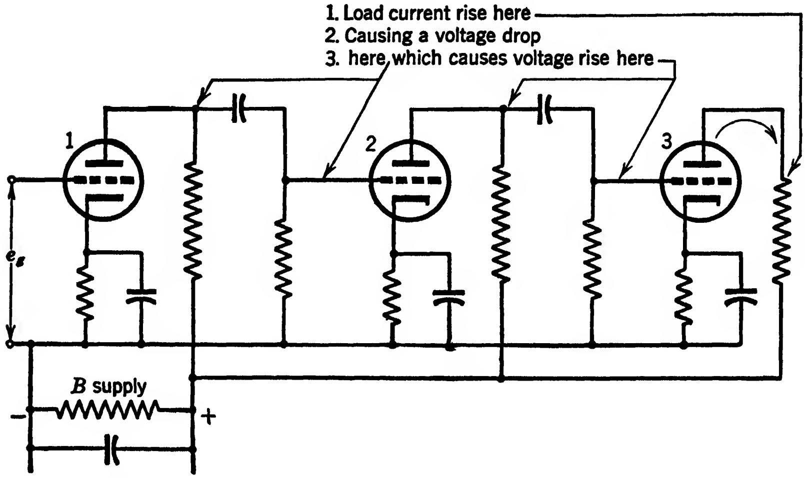


Fig. 11. Circuit illustrating the cause of "motor boating."

having appreciable internal resistance is used to feed several stages and when a low-frequency signal is being amplified. As an example, consider a three-stage resistance-coupled amplifier having a common $B+$ supply, as shown in Fig. 11. If for some reason the plate current in the output stage 3 increases, the $B+$ voltage will drop slightly because of the internal resistance of the supply and the impedance of the supply (power pack) at low frequencies. This change in voltage is impressed on the grid of the second stage 2 through the plate load resistance of the first. Since a decrease in voltage on grid 2 will be amplified and appear as an increase at plate 2, the grid of the third stage will be driven more positive, thus increasing the plate current and further reducing the common $B+$ voltage. The action of the second stage on the $B+$ supply is exactly opposite that of the third. However, because of the gain through the amplifier, the third stage will have the largest change in plate current, and hence

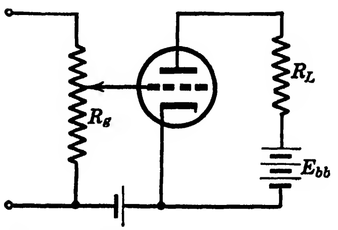
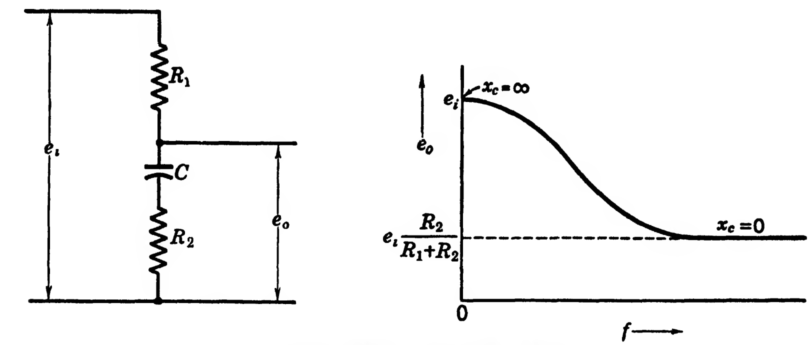
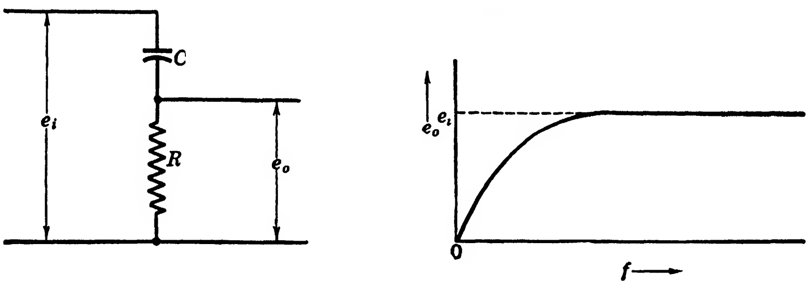


Fig. 12. Simple circuit for volume control.

the major effect on the $B+$ supply voltage. The increase in plate current on the third stage continues until this tube reaches saturation. At that time the action reverses and the final grid is soon driven to cutoff. This action takes place at a very low frequency, usually only a few cycles per second, determined by the various RC combinations in the



(*a*) Bass boost or treble attenuation



(*b*) Bass attenuation

FIG. 13. Circuits for tone control.

circuit. It can be prevented by using a power supply with extremely low internal resistance, or by "decoupling" the $B+$ supply for each stage through a series resistor and shunt condenser.

Amplifier gain may be controlled by varying the bias on a stage which uses a variable-mu screen-grid tube (see page 66). A more common method employs a potentiometer in place of one of the grid resistors with the grid connected to the movable tap (Fig. 12). It is usually undesirable to place a potentiometer directly in the input of a high-gain amplifier because of noise generated when the control is moved. However, the amplification ahead of the gain control should be limited so that a large signal cannot cause overloading.

Tone controls are used to vary the frequency response of the amplifier. With the simple *RC* circuits (Fig. 13) it is possible to obtain bass boost, bass attenuation, or treble attenuation. High-frequency boost is unsatisfactory because it accentuates harmonic distortion.

Amplifier Feedback Circuits. Feedback in an amplifier is the transfer of power from the plate output circuit back into the grid input

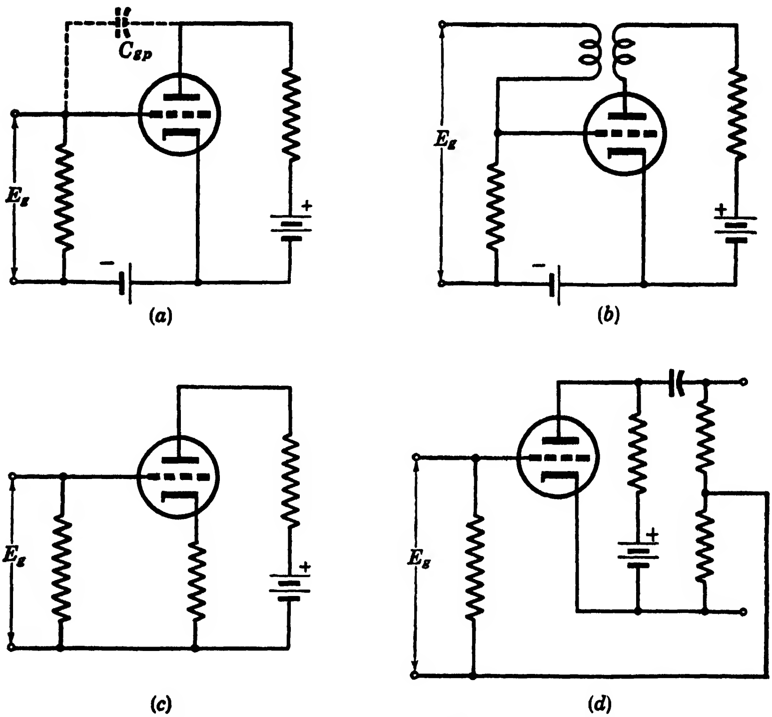


FIG. 14. Forms of feedback coupling circuits in amplifiers.

circuit. Generally, the signal fed back has approximately the same wave form as the input signal, but its phase may vary from that of the input signal. Usually the signal fed back will be either in phase with or 180 degrees out of phase with the input. The former phase relation is called *positive or regenerative feedback,* the latter is known as *negative or degenerative feedback.* With positive feedback the gain of the amplifier is increased; with negative feedback the gain will be reduced. Feedback may be intentional or incidental in the operation of an amplifier circuit but it is always an important and sometimes a very useful phenomenon.

Feedback in an amplifier is a result of a coupling between the output and the input circuits. Such coupling may be capacitive (electric field), transformer (magnetic field), or some form of direct coupling where an impedance is common to both the output and the input circuits. Four forms of feedback coupling circuits are illustrated in Fig. 14. In part *a* a capacitive coupling between the plate and the grid is suggested. This coupling may be due to the interelectrode capacity of the tube itself or a capacitor external to the tube. Such a circuit may set up undesirable oscillations in a high-frequency circuit, or it may serve to reduce the input impedance to the tube as discussed earlier. Part *b* of Fig. 14 shows a transformer coupling. This form of amplifier circuit known as a "tickler" connection was used as a regenerative circuit on some early radio receivers to give high gain with only one tube. This circuit was unstable and served to rebroadcast radiation so that, in general, it has been abandoned since the early days of amateur radio. A modification of this circuit is sometimes used in oscillators and will be discussed in the following chapter. The circuit of part *c*, Fig. 14, is classed as current feedback since the signal is fed back into the cathode resistor through current variations in the plate circuit. This circuit and action forms the basis of the cathode follower discussed in the preceding chapter. Part *d* of Fig. 14 is a voltage feedback circuit in which a portion of the output voltage is fed back to the grid in series with the incoming signal.

Negative Feedback. Negative feedback in an amplifier serves to stabilize (hold nearly constant) the gain in an amplifier and also

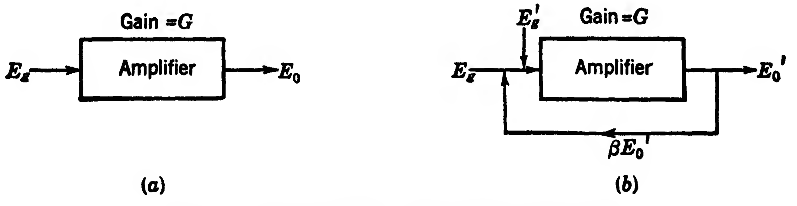


FIG. 15. Block diagram of negative feedback.

serves to reduce distortion and noise generated within the amplifier. These advantages can be understood by a study of the block diagrams of Fig. 15 and subsequent developments and problems. In part *a* of the figure a simple amplifier without feedback is indicated. By definition the gain G is the ratio of the output voltage E_0 to the input signal voltage E_g. With the addition of the feedback circuit of part *b*, Fig. 15, a portion of the output β is fed back to input where β is defined by the ratio

$$\beta = \frac{\text{voltage fed back}}{\text{output voltage}}$$

For this circuit the new input voltage E_g' for the amplifier is

$$E_g' = E_g + \beta E_0'$$

and

$$E_0' = GE_g'$$

Therefore

$$E_g' = E_g + G\beta E_g'$$

$$= \frac{E_g}{1 - G\beta} \tag{9}$$

$$E_0' = \frac{GE_g}{1 - G\beta} \tag{10}$$

and the gain G' with feedback is

$$G' = \frac{\dfrac{GE_g}{1 - G\beta}}{E_g} = \frac{G}{1 - G\beta} \tag{11}$$

Since G and β are voltage or current ratios, they have both magnitude and phase angle. When the product of these ratios $G \times \beta$ is unity and the phase angle is zero, the quantity $(1 - G\beta)$ becomes zero, making the output infinitely large. Here instability is extreme and would cause the amplifier to oscillate or "sing" around the closed (feedback) loop.

The stability of the amplifier can be understood by solving a numerical example. For the amplifier of Fig. 15, let the gain $G = 80$ db (voltage ratio of input to output of 1 to 10,000) and the loss in the feedback circuit be 60 db (voltage ratio of 1000 to 1). If the input voltage E_g is 1 millivolt, the actual input voltage to the amplifying unit E_g' is, from equation 9,

$$E_g' = \frac{1}{1 - 10{,}000/-1000} = \frac{1}{1 - (-10)} = \frac{1}{11} = 0.09091 \text{ millivolt}$$

The output voltage is

$$G \times E_g' = 10{,}000 \times 0.09091 = 909.1 \text{ millivolts}$$

This output of 909.1 millivolts is also impressed on the feedback circuit which allows $\frac{1}{1000}$ of it to be fed back to the input circuit. In the feedback circuit the phase is shifted so that the feedback voltage has a

minus sign. Hence -0.9091 millivolt combines with the initial 1 millivolt to form the actual input to the amplifier, which gives

$$1.000 - 0.9091 = 0.0909 \text{ millivolt}$$

This result checks the value of E_g' obtained above, which means that the amplifier is stable and that, as long as the applied 1 millivolt is maintained, there will be 909.1 millivolts in the output. The over-all gain of the amplifier under these conditions is

$$20 \log_{10} \frac{\text{output voltage}}{\text{input voltage}} = 20 \log_{10} \frac{909.1}{1}$$

$$= 20 \times 2.9586 = 59.17 \text{ db}$$

It should be noted that the over-all gain is practically the same as the loss β in the feedback circuit.

Now if we increase the gain of the amplifier to 100 db (voltage ratio of input to output of 1 to 1,000,000) and use the same loss in the feedback circuit, a new set of computations will be as follows:

$$E_g' = \frac{1}{1 - 100{,}000/-1000} = \frac{1}{101} = 0.009901 \text{ millivolt}$$

The output voltage $G \times E_g'$ will be

$$100{,}000 \times 0.009901 = 990.1 \text{ millivolts}$$

and the over-all gain of the amplifier is

$$20 \log_{10} \frac{990.1}{1} = 20 \times 2.9957 = 59.91 \text{ db}$$

Thus the over-all gain is substantially the same as before and nearly equal to the loss in the feedback circuit. This unexpected result is explained by the fact that, as the amplification is increased, the feedback circuit feeds back a larger voltage 180 degrees out of phase, which combines with the applied input voltage to form a lower actual input to the amplifier, thereby reducing the output voltage to the point where the over-all gain is practically the same as the loss in the feedback circuit. The excellent stability and constancy of gain of the negative feedback amplifier cause it to be used rather widely in telephone repeaters, radio circuits, and equipment for electrical measurements.

Some distortion will occur in a final power stage of an amplifier. The larger the gain, the greater the voltage swings will be over the

nonlinear dynamic characteristic curve with corresponding increase in distortion. If negative feedback is added to the final stage the over-all gain will be reduced and *the distortion will be reduced in the same proportion.* However, the loss in gain in this stage can be compensated by an increase in gain in the preceding stages of voltage amplifiers which have low distortion. Thus the resulting distortion in the output of the complete amplifier group may be reduced by negative feedback.

Negative voltage feedback also reduces the effective plate resistance making it especially helpful with pentodes and beam-power tubes which must work into a load (such as loudspeaker) that varies with frequency. Where it is desirable to maintain constant output current, a voltage proportional to this current is fed back to the grid circuit. Since the feedback voltage must be 180 degrees out of phase with the signal at all frequencies, it is essential that the phase shift of all intervening circuits must be kept small. The feedback loop may include only the output stage or may be extended around the driver and one or more voltage amplifier stages.

PROBLEMS

1. Use the characteristic curves of a 6L6 (triode connection) and calculate the power output and per cent second harmonic distortion for the following loads: (a) 2000 ohms; (b) 3500 ohms; (c) 5000 ohms. Assume $E_{bb} = 250$ volts, $E_{cc} = -20$ volts, $E_{gm} = 20$ volts peak, static load resistance zero.

2. A 6L6 (tetrode connection) is used as a Class A power amplifier. If $E_{bb} = 250$ volts, $E_{c2} = 250$ volts, $E_{cc} = -14$ volts, $E_{gm} = 14$ volts peak, and $R_L = 2500$ ohms, calculate: (a) the per cent second harmonic distortion; (b) the per cent third harmonic distortion. (c) If higher order distortion is neglected, calculate the total harmonic distortion ($D_t = \sqrt{D_2^2 + D_3^2}$); (d) the rms power output.

3. Repeat Problem 2 for a 6F6 (pentode connection) with $E_{bb} = 250$ volts, $E_{sc} = 250$ volts, $E_{cc} = -16.5$ volts, $E_{gm} = 16.5$ volts peak, $R_L = 7000$ ohms.

4. Using data from a tube manual, design a single-ended Class A amplifier using a type 6V6 tube with a plate supply voltage of 250 volts. The tube is to work into a 600-ohm load. Specify the turns ratio of the output transformer and the minimum primary inductance if the gain is to drop to 70.7% of the middle-range gain at 75 cps. If cathode bias is used, what values should R_k and C_k have?

5. A pair of type 2A3 tubes is used in a push-pull Class AB_1 amplifier. If the amplifier is operated as shown in Fig. 4, (a) calculate the rms power output. (b) What turns ratio should the output transformer have if the load R is 500

ohms? (c) If the input transformer has a 1:3 turns ratio, what input voltage is necessary for maximum power output?

6. Two type 6F6 tubes (triode connection) are used in a push-pull Class AB_2 amplifier. Assume $E_{bb} = 350$ volts, $E_{cc} = -38$ volts, E_g (grid-to-grid peak) = 123 volts, R_L (plate-to-plate) = 6000 ohms. Draw the composite characteristics and repeat Problem 5.

7. A pair of type 46 tubes is used in a push-pull Class B amplifier. The amplifier is to work into a 500-ohm load.

$I_{p\,max} = 170$ ma	$E_{bb} = 300$ volts
$e_{b\,min} = 80$ volts	I_b (max signal) = 115 ma

(a) What turns ratio should the output transformer have? (b) Calculate the rms power output. (c) What is the maximum signal efficiency?

8. A type 304TH tube is used as a Class C rf amplifier.

$E_{bb} = 1500$ volts
Maximum plate dissipation = 300 watts
Assume an efficiency of 70%

(a) Calculate the maximum permissible d-c plate current. (b) Calculate the rf power output.

9. In Fig. 16 is shown an amplifier with inverse feedback. The gain of the amplifier *without* inverse feedback is 25. $R_1 = 82{,}000$ ohms; $R_2 = 8000$ ohms. (a) What is the value of β? (b) What is the gain with inverse feedback?

10. In Fig. 16 a 6L6 is used in a circuit employing inverse feedback to reduce distortion. The amplifier is operated Class A with $E_{bb} = 250$ volts, $E_{cc} = -14$ volts, $R_L = 2500$ ohms. (a) Calculate the maximum power output of the amplifier without inverse feedback. (b) If the magnitude of β is 0.1, what is the power output? (c) If $R_1 = 50{,}000$ ohms, what value should R_2 have? (d) What value should E_g have to obtain maximum power output with inverse feedback?

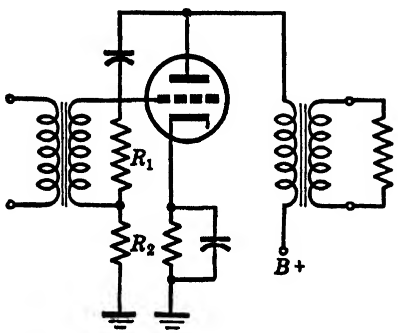


FIG. 16

Chapter IX

ELECTRON-TUBE OSCILLATORS

Electron-Tube Oscillators. An electron-tube oscillator is a combination of electron tubes and circuits which serves to convert direct current into some form of periodic varying current. The resulting current may be sinusoidal with the oscillator functioning as an inverter, or the wave form may be nonsinusoidal having a square, sawtooth, or pulse shape. Many oscillators operate by utilizing the

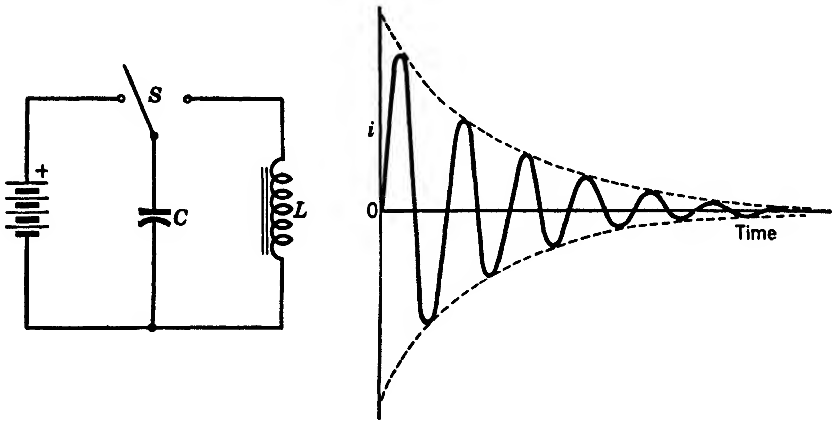


FIG. 1. Simple oscillator circuit.

amplifying power of vacuum tubes and the resonating property of tuned LC circuits.

Oscillations occur naturally in the simple parallel LC circuit shown in Fig. 1. If switch S is thrown to the left the battery will charge the capacitor by removing electrons from the top plate and storing them on the lower plate. This action stores energy in the electric field of the capacitor. Now turn the switch S to the right and the capacitor will discharge through the inductance L with electrons moving from the lower plate back to the upper. The rising current through L will store energy in the magnetic field surrounding it. When C becomes discharged, the energy of its charge will have been transferred to the

magnetic field of L. This stored energy in L will continue the flow of electrons and begin to charge C with a reversed polarity. This process continues until all the energy in the magnetic field has been transferred to C. At this point C begins to discharge again with a reversed direction of electron flow. Obviously, when C has released all its stored energy to the inductance L, the latter will have acquired energy to recharge C with the same polarity as originally provided by the battery. Now the circuit is restored to its original condition and is ready to repeat the process. If both the inductance and capacitance were without resistance or any form of loss, the resulting ideal circuit would continue to oscillate indefinitely. Ideal circuits cannot be realized and some resistance is always present. Such resistance will reduce each swing of current as illustrated in the right view of Fig. 1. The larger the value of circuit resistance, the more rapidly the oscillations will be damped out. A freely swinging pendulum will have its oscillations damped out with time (like right view of Fig. 1). In a clock the pendulum is kept swinging with a uniform stroke by adding enough mechanical energy to each stroke to supply the losses due to friction and windage. In a similar manner, the LC circuit of Fig. 1 may be made to continue oscillations of uniform magnitude by adding the necessary electrical impulse at each swing. The LC circuit having a high Q has a natural period or frequency f_r which is determined by the equation,

$$f_r = \frac{1}{2\pi\sqrt{LC}} \quad \text{(equation 3, Chapter V)}$$

and the magnitude of this frequency can be controlled through changes of L and C.

Electronic oscillators may be classified on the basis of (1) wave form produced, or (2) the principle employed for excitation. The output wave form may be sinusoidal or nonsinusoidal, and the excitation may be self-generated or external. The four common methods of excitation or control which constitute a basis for oscillator classification are: (*a*) feedback, (*b*) negative resistance, (*c*) mechanical, (*d*) relaxation.

The vacuum-tube oscillator is employed for the generation of alternating current at frequencies beyond the range of the rotary type of alternator. Rotating alternators operate efficiently and deliver large amounts of energy at 60 cycles, but their principle of operation limits design to a few thousand cycles at reduced output capacities. The vacuum-tube oscillator does not have such power and frequency limi-

tations but is capable of outputs ranging from a few microwatts to over a million watts at frequencies from a few cycles per second to thousands of megacycles.

Feedback Oscillators. A feedback oscillator is a self-excited amplifier. In the preceding chapter it was suggested how positive or regenerative feedback could be employed to cause an amplifier to produce sustained oscillations. Such feedback may be effected by magnetic (transformer) coupling, electric field (capacitive) coupling, or a direct voltage coupling between the plate and the grid circuits. All these methods are employed in practice and are illustrated in the circuits

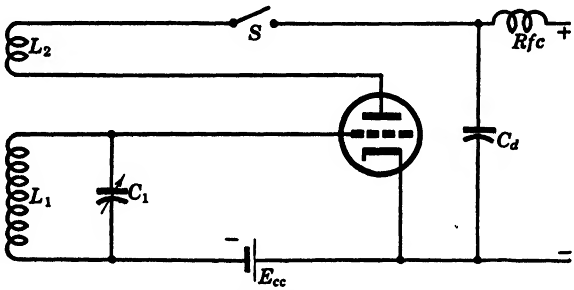


FIG. 2. Simple feedback oscillator circuit.

of Figs. 2 and 3. A simple circuit illustrating magnetic coupling is given in Fig. 2. In this circuit the feedback or "tickler" coil L_2 in the plate output is coupled inductively to coil L_1 of the tuned L_1C_1 circuit, which, in turn, establishes the signal voltage applied to the grid. When the switch S in the plate circuit is closed, current flows in the cathode-plate circuit through L_2. The rising flux in L_2 threads coil L_1 inducing a voltage which charges C_1. When the plate current reaches its normal value, the energy in the magnetic field of L_1 overruns, charging C_1 to a higher potential. After reaching a peak level, C_1 discharges into L_1 and the L_1C_1 tuned circuit oscillates at a frequency determined by its resonant frequency. The oscillating grid signal causes the plate current to oscillate, which, in turn, feeds back enough energy to overcome the losses in the L_1C_1 tuned circuit. Thus the entire circuit of Fig. 2 becomes a self-excited amplifier.

The conditions necessary for sustained oscillations in a self-excited vacuum-tube oscillator are:

1. The feedback voltage from the plate circuit to the grid must have a phase reversal of 180 degrees.

2. The power fed back must be sufficient to supply the losses in the grid input.

3. A tuned (LC) or other frequency selective circuit must be used to establish an oscillating frequency.

Grid bias for an oscillator may be obtained from a fixed source if Class A or Class B operation is employed. However, the need for high efficiency requires the use of Class C operation with a bias approximating two times cutoff value. Under this condition plate current cannot flow and oscillations will never start unless the bias is reduced momentarily. Accordingly, an RC biasing circuit or component is standard on many oscillators. This RC component performs two important functions. First, it serves to start oscillations, and, second, it aids in maintaining stability in the amplitude of the oscillations produced in the plate circuit. When switch S in Fig. 3a is closed, a surge of plate current flows since the grid initially is at approximately zero potential. The rise of plate current induces a voltage in coil L_1, which tends to make the grid potential move in the positive direction. With the rise in plate current more electrons arrive at the grid, which tends to increase the grid bias. At the same time the voltage across L starts oscillations in the LC tuned circuit at its resonant frequency. These oscillations vary the potential between the grid and the cathode, which, in turn, vary the plate current. These variations in plate current feed energy back through the coupling into the grid circuit. The first oscillation will be small in magnitude, but succeeding oscillations will increase rapidly in magnitude until the energy generated in the plate circuit equals the losses in the plate circuit, grid circuit, and the output circuit. The *grid circuit is nonconducting* on the negative half-cycle of operation, and the condenser C_g discharges slowly through R_g. With *suitable* values of R_g and C_g this combination will maintain a satisfactory minimum negative potential. Thus, if for any reason the amplitude of plate current swings rises, the potential across capacitor C_g rises and a more negative grid results, whereas, if the amplitude of the plate current swings decreases, the grid tends to become less negative and permits more plate current to pass. In this manner the action of the R_gC_g component promotes stability in the amplitude of plate output. In passing it should be noted that, if the magnitude of the time constant of the R_gC_g component should be too large, the discharge of C_g through R_g might become too slow so that the negative grid might stop oscillations momentarily. This condition would cause an intermittent instead of a steady operation of the oscillator.

All circuits using the grid leak RC combination are theoretically self-starting. However, it is possible that under certain circuit con-

ditions oscillations will not start. If this happens, the bias will remain near zero and the tube will draw a large current from the $B+$ supply

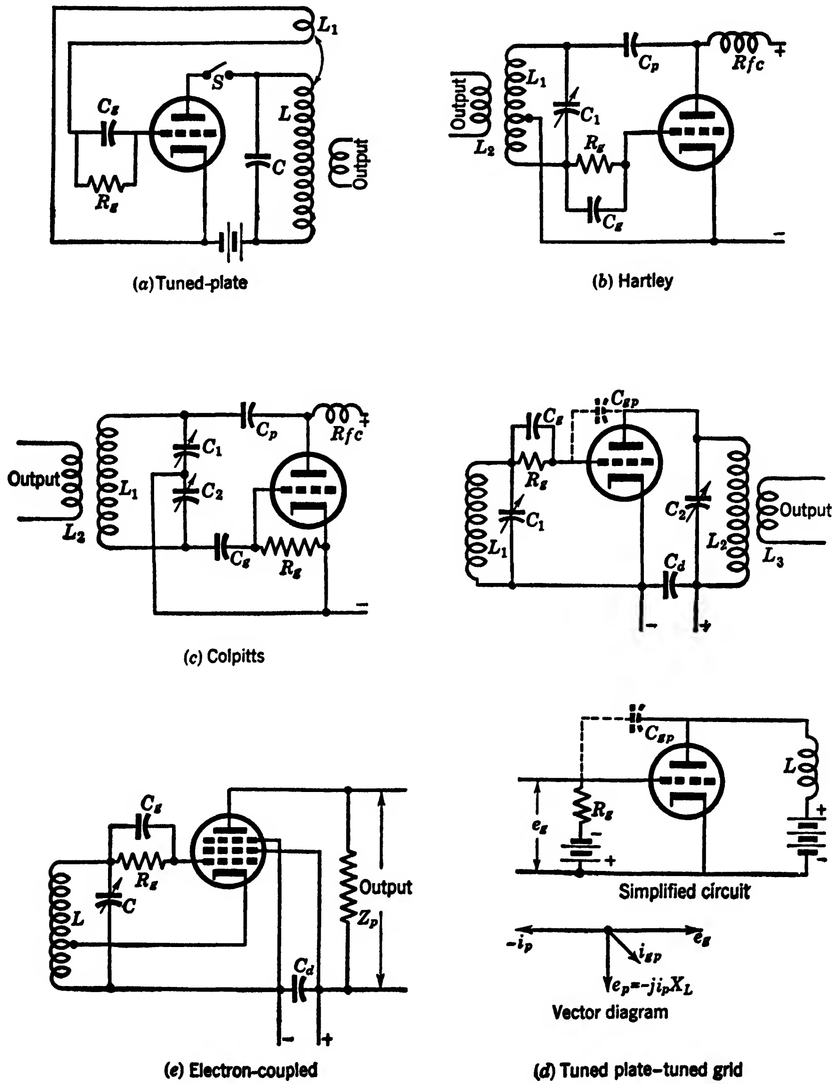


FIG. 3. Tuned-oscillator circuits.

and is likely to overheat. For protection against this contingency, a portion of the bias may be obtained from a cathode resistor or from a fixed source. The values of grid leak and grid capacitor are not criti-

cal. In practice, the magnitude of the resistor should be adjusted to provide the proper bias for maximum output. In Fig. 2, C_d is a by-pass condenser and the radio-frequency choke *Rfc* has a high reactance to the operating frequency, thus serving as a filter to prevent the a-c output current from flowing through the power supply.

The operating frequency and the stability of that frequency for a self-excited oscillator is determined primarily by the tuned LC circuit. The resonant frequency for an ideal parallel LC circuit is

$$f_r = \frac{1}{2\pi\sqrt{LC}}$$

In the physical circuit of an oscillator the coil L contains some ohmic resistance, and, in addition, the impedance of the output load is reflected into L so as to raise the effective resistance and reactance in the LC loop (Fig. 3*a*). The result of such effective resistance is to cause the circuit to oscillate at a frequency differing from the resonant frequency f_r given in the preceding formula. The manner of this frequency variation will be discussed in an analysis given later in this chapter. It follows that changes in temperature of coil L and changes in load impedance will change the effective resistance in the LC circuit, and thus vary the operating frequency and frequency stability of an oscillator. Variations in frequency can be reduced to small values by using coils having a high Q and by using a high ratio of C to L.

Sinusoidal output from a self-excited oscillator will exist if sufficient energy is stored during the interval of plate current flow to supply the load until the next plate current pulse. The storehouse or "tank" for this energy is the tuned circuit. An excellent mechanical analogy to the tank circuit is the 4-cycle gasoline engine. During the combustion stroke of the piston kinetic energy is stored in the rotating flywheel. For the next three strokes (exhaust, intake, and compression), the flywheel must supply all the losses of the engine and the load. As kinetic energy is lost, the angular velocity of the flywheel decreases. Hence, the speed between explosions is not constant, and the greater the load for a given flywheel, the more pronounced will be the variation. In the vacuum-tube circuit the function of the flywheel is performed by the tuned tank which causes the current to oscillate, flowing to the condenser on one half of the cycle and to the inductance on the succeeding half-cycle. It should be noted that the tank current is a true alternating current and generally is many times

the magnitude of the d-c plate current. This is an important consideration in selecting the circuit components since, with tubes of small capacity, tank currents of several amperes may flow. A larger capacitor and a smaller inductance increase the circulating current, and act in the same way as increasing the inertia of the flywheel. High C to L ratio maintains a more nearly constant rate of oscillation between plate current pulses and results in lower harmonic distortion. It has been found that for stable operation the ratio of energy stored to energy lost per cycle should be not less than 2, and the circulating volt-amperes should be at least 4π times the power output.* These conditions are met when the effective Q of the tank is 4π or greater. Q can be readily adjusted by changing the load coupling.

Frequency stability of an oscillator circuit may be defined as the per cent variation of frequency output from the mean frequency arising from fluctuations in circuit parameters, supply voltages, and other factors.

The frequency stability which can be obtained with an LC resonant circuit varies approximately from 0.001 per cent to 0.01 per cent per degree C. Vibration, temperature, and humidity have the greatest effect on frequency stability but may be neutralized by proper construction of the coil and condenser. Temperature effects in the coil may be corrected by selecting a condenser having an equal and opposite temperature coefficient. Plate and filament voltages also influence frequency unless their effects are balanced by a "stabilized" circuit. If frequency stability is important the design of a self-excited oscillator should provide:

(*a*) High Q, a high ratio of C to L, and very light loading.

(*b*) Rugged construction to eliminate vibration of parts and wiring.

(*c*) Temperature compensation.

(*d*) Humidity protection.

(*e*) Regulated plate and filament voltages or some form of stabilization.

Under favorable conditions the over-all frequency stability may be in the order of 0.005 to 0.01 per cent.

There are numerous oscillator circuits that employ a combination of the RC grid leak and the tuned LC components. Five useful combinations are illustrated in Fig. 3. Part *a* is a tuned-plate circuit which will be analyzed in some detail subsequently. A widely used circuit, known as the Hartley, is given in Fig. 3*b*. In this circuit excitation is

* Glasgow, *Principles of Radio Engineering*, McGraw-Hill, p. 272.

adjusted by moving the cathode tap on the tank coil. Normally about one-third of the total turns are required for the grid section. This circuit has the minimum of component parts, and the frequency can be adjusted easily without much effect on the grid excitation.

In the Colpitts circuit, Fig. 3*c*, the cathode tap is made in the capacitive branch of the tuned circuit. Capacitors C_1 and C_2 are adjusted to provide the proper grid voltage and at the same time to give the desired operating frequency. This circuit may be somewhat more difficult to adjust than the Hartley since the two variable condensers must be controlled together when frequency is changed. The Colpitts circuit is preferable for high-frequencies since the grid voltage leads the plate voltage by an amount which tends to correct for electron transit time. In other respects, the Colpitts circuit is similar to the Hartley.

An oscillator using tuned *LC* circuits for both the plate and the grid is illustrated in Fig. 3*d*. Feedback voltage is obtained through the grid-plate capacitance of the tube, or, if this is insufficient, through a small external capacitor. The operation of this circuit is as follows. If the plate load consists of an inductance whose reactance is small compared to r_p, the plate voltage will lead the current by 90 degrees, and hence will lag the grid voltage e_g by 90 degrees instead of 180 degrees (Fig. 3*d*, lower). Also, this voltage considered across C_{gp} and R_g in series causes a current i_{gp} which leads e_p because of the capacity reactance. From this vector diagram it is apparent that i_{gp} has a component in phase with e_g which will add to any signal current. In the case of the tuned plate-tuned grid oscillator, R_g represents the effective resistance of the grid tank and input capacitance of the tube when this circuit is tuned to the exact operating frequency. Load L is the effective impedance of the plate tank, which must be adjusted to appear as an inductive reactance at the operating frequency. This factor is important because regenerative feedback cannot exist unless the plate load is inductive. In practice, L must also contain a resistive component to represent the losses in the plate circuit and the load. The operating frequency of this type of oscillator is determined primarily by the grid tank, although the plate circuit must be adjusted also within fairly close limits. Tubes having extremely low grid-to-plate capacity, such as those with screen grids, will not oscillate in this circuit unless a small external capacitor is connected between the plate and the grid.

An *electron-coupled* oscillator circuit is illustrated in Fig. 3*e*. This circuit uses a screen grid or pentode connected so that its cathode, con-

trol grid, and screen grid act as an oscillator similar to the Hartley circuit of part *b*, while the plate circuit of the tetrode serves in the capacity of an amplifier. The coupling between the two circuits and the two functions is the electron stream within the tube, and hence the term electron-coupled oscillator. Since the screen grid is held at rf ground potential and also serves as a shield between the two circuits, this oscillator is very stable because load variations have little effect on the frequency. Another factor which aids the stability of the electron-coupled oscillator is that an increase in d-c screen voltage will decrease the frequency, while an increase of d-c plate voltage will increase the frequency. Accordingly, variations in the power supply voltage have little effect on the frequency of oscillation.

Two types of power-supply feed to the oscillator, series and shunt, are illustrated in Fig. 3. For the Hartley and Colpitts circuits the power supply is in "shunt" with cathode-plate circuit of the tube and in shunt with the tank and the grid circuits. This arrangement plus the use of the coupling capacitor C_p serves to isolate the high-voltage d-c potential from the grid and the tank circuits. The capacitor C_p must be large enough to carry the a-c plate output. The choke Rfc in the $B+$ lead of the power supply must have a high value of reactance at the operating frequency since it is effectively in parallel with the L of the tank circuit. Series feed of the power supply to the oscillator is illustrated in parts *a, d,* and *e* of Fig. 3 where the plate current from the power supply flows through some portions of the feedback or tank circuit.

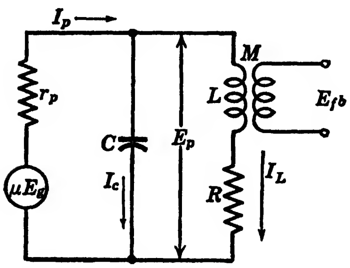


FIG. 4. Equivalent circuit of a tuned-plate oscillator.

Analysis of a Tuned-Plate Oscillator. An analysis of the operation of an oscillator circuit may be made by the methods employed on amplifiers in the preceding chapters. The tuned-plate circuit of Fig. 3*a* may be reduced to the equivalent circuit of Fig. 4, and Class A operation with zero grid current will be assumed for the purpose of the study though it should be understood that Class C operation is frequently involved in tuned-plate circuits.

An analysis may begin with the equation for the gain of an amplifier as developed for the equivalent circuit.

$$\text{Gain} = \frac{E_p}{E_g} = -\frac{\mu Z_L}{r_p + Z_L} \quad \text{(equation 7, Chapter VI)}$$

In order to sustain oscillations in the tuned-plate oscillator the feedback voltage E_{fb} of Fig. 4 must be equal to E_g in magnitude and phase. Thus the ratio

$$\frac{E_p}{E_{fb}} = \frac{E_p}{E_g} = -\frac{\mu Z_L}{r_p + Z_L} = \frac{-1}{\dfrac{r_p + Z_L}{\mu Z_L}} = \frac{-1}{\dfrac{1}{g_m Z_L} + \dfrac{1}{\mu}} \quad \text{for sustained (oscillation)} \tag{1}$$

Again, from Fig. 4, it follows that *

$$\frac{E_p}{E_{fb}} = \frac{-(R + j\omega L)I_L}{j\omega M I_L} = -\frac{R + j\omega L}{j\omega M} \tag{2}$$

and

$$Z_L = \frac{(R + j\omega L)\dfrac{1}{j\omega C}}{R + j\left(\omega L - \dfrac{1}{\omega C}\right)} \tag{3}$$

The substitution of the expressions of equations 2 and 3 into equation 1 gives

$$\frac{R + j\omega L}{j\omega M} = \frac{1}{\dfrac{1}{\mu} + \dfrac{R + j\left(\omega L - \dfrac{1}{\omega C}\right)}{g_m(R + j\omega L)\dfrac{1}{j\omega C}}}$$

from which

$$\left(\frac{R}{g_m} + \frac{L}{\mu C} - \frac{M}{C}\right) + j\left[\frac{1}{g_m}\left(\omega L - \frac{1}{\omega C}\right) - \frac{R}{\mu\omega C}\right] = 0 \tag{4}$$

If equation 4 holds true, both the real and imaginary components must be equal to zero. Equating these components to zero and solving the resulting equations gives some useful expressions covering the criteria for sustained oscillations.

$$\frac{R}{g_m} + \frac{L}{\mu C} - \frac{M}{C} = 0 \tag{5}$$

$$\frac{1}{g_m}\left(\omega L - \frac{1}{\omega C}\right) - \frac{R}{\mu\omega C} = 0 \tag{6}$$

* For a more complete treatment of this subject the reader is referred to *Applied Electronics*, M.I.T. Electrical Engineering Staff, John Wiley & Sons, pp. 599–608, from which this treatment has been taken with permission.

From (5)

$$g_m = \frac{R}{\frac{M}{C} - \frac{L}{\mu C}} = \frac{\mu RC}{\mu M - L} \tag{7}$$

Equation 7 shows the value of the transconductance of the tube necessary for sustained oscillations for specified values of μ, R, C, L, and M. From equation 6

$$\omega^2 = \frac{\mu + Rg_m}{\mu LC} = \frac{1}{LC}\left(1 + \frac{Rg_m}{\mu}\right) \tag{8}$$

Since for the parallel resonant circuit, ${\omega_0}^2 = 1/LC$, equation 8 reduces to

$$\omega^2 = {\omega_0}^2\left(1 + \frac{Rg_m}{\mu}\right) \tag{9}$$

and

$$\omega = \omega_0\sqrt{1 + \frac{Rg_m}{\mu}} = \omega_0\sqrt{1 + \frac{R}{r_p}} \tag{10}$$

Equation 9 shows that the frequency of oscillation is slightly higher than the resonant frequency of the tuned-plate load. However, for coils with a high Q, the resistance R of the tuned-plate circuit is low, and hence the frequency of oscillation approaches closely the frequency of the tuned circuit.

A substitution of equation 7 into equation 9 gives

$$\omega^2 = \frac{1}{LC}\left(1 + \frac{R}{\mu} \cdot \frac{R\mu C}{\mu M - L}\right) = \frac{1}{LC}\left(\frac{1 + R^2C}{\mu M - L}\right) \tag{11}$$

which shows how the frequency of oscillation is controlled by the constants of the tuned-plate circuit.

Crystal-Controlled Oscillators. The frequency of oscillations generated by the tickler coil, Hartley, and Colpitts oscillator circuits is affected considerably by changes in load, supply voltages, and temperature. Although the variation in frequency is small in electron-coupled oscillators, it is sufficient to be objectionable in broadcast transmitters, telephone carrier systems, and similar applications. Where a precision control of the frequency is necessary, crystal-controlled oscillator circuits are employed.

Certain crystalline substances, such as quartz, Rochelle salts, and tourmaline, exhibit mechanical and electrical properties known as the piezoelectric effect. Thus, if a mechanical force is applied to one of

these substances, a voltage is developed. Conversely, if a thin slab of the substance is connected to a source of alternating voltage, it changes its physical shape and produces mechanical vibrations. Thin slabs cut from quartz crystals for use in crystal oscillator circuits are called *crystals*. When such a crystal is vibrating at its resonant frequency, it requires only a small mechanical force of the same frequency to obtain vibrations of a large amplitude. The mechanical resonant frequency of a crystal depends chiefly on its thickness. When the frequency of an alternating voltage applied to a crystal is the same as its resonant (mechanical) frequency the crystal will vibrate, and only a small voltage is necessary to keep it vibrating. In a similar manner, a crystal set in mechanical vibration will generate a relatively large voltage at its resonant frequency. If this crystal is placed between the grid and the cathode of a vacuum tube and a small amount of energy is taken from the plate circuit and applied to the crystal to keep it vibrating, the circuit will act as an oscillator. The natural frequency of a crystal is critical (and *precise*). Thus if the constants of the oscillator circuits are properly adjusted, the crystal will assure a *precise* frequency output.

Research and application have developed the art of producing and applying quartz crystals in many electronic circuits. If a quartz (mother) crystal is cut parallel to one face of the crystal, the resulting slab is said to be Y-cut and its frequency will be about 2×10^6 cycles per millimeter of thickness. Frequency will increase with temperature at the rate of 25 to 100 parts per million per degree C. X-cut crystals (cut perpendicularly to a face) are slightly thicker for the same frequency and have a negative temperature coefficient of 10 to 25 parts per million per degree C. Other cuts can be made to give zero temperature coefficients, such as AT and CT crystals, which vibrate in shear. The latter are oriented so that their frequency is determined by the large dimension of the plate and are desirable, therefore, for lower frequencies.*

Crystals are usually mounted between two flat metal plates (Fig. 5) although sometimes the electrodes are plated directly on the quartz. From the equivalent circuit of Fig. 5*b* it can be seen that the crystal acts like a resonant *LC* circuit and can be used in place of the grid tank in a tuned plate-tuned grid oscillator, as in Fig. 6. Excessive

* S. C. Hight, and G. W. Willard, "A Simplified Circuit for Frequency Substandards Employing a New Type of Low-Frequency Zero-Temperature-Coefficient Quartz Crystal," *Proceedings I.R.E.*, Vol. 25, p. 549, May 1937.

feedback must be avoided to prevent the mechanical vibrations from fracturing the crystal.

The equivalent Q of a good crystal is extremely high, reaching in some cases a value of 500,000. Thus the over-all frequency variation in a crystal oscillator may be made less than 1 part in a million. Quartz crystals can be produced to operate at frequencies as low as

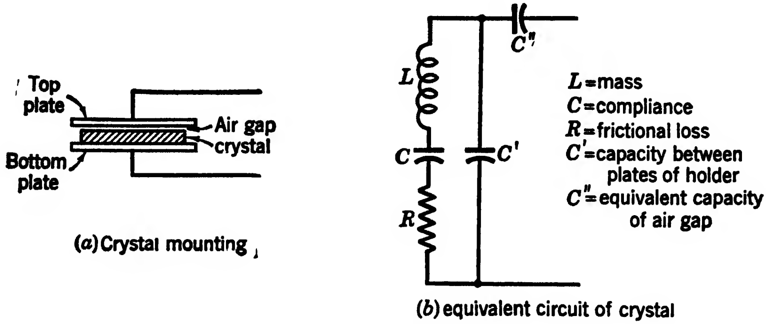


FIG. 5

5000 cycles though the usual range is 50 kilocycles to 50 megacycles. They are widely used to control the frequency of radio transmitters, of frequency standards, and of the oscillator in superheterodyne receivers designed to operate on fixed frequencies. The resonant properties and extremely high Q of quartz crystals may also be used as elements in filter circuits designed to pass or reject a very narrow band of frequencies. Such filters have proved useful in wave analyzers and

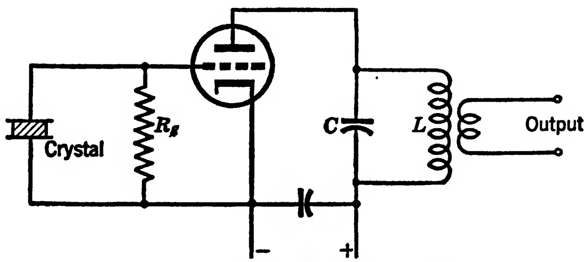


FIG. 6. Crystal oscillator circuit.

in communication receivers where it is desired to eliminate interference from other stations or nearby channels. Crystal filters are widely used in carrier telephone circuits.

***RC* Oscillators.** The oscillators that have been considered have utilized resonance in some form, either in *LC* tuned circuit or through the piezoelectric properties of quartz. Another group of feedback

oscillators uses capacitors and resistors as the coupling element and operates by virtue of the phase shift produced by *RC* circuits. An amplifier circuit having its output connected back to the input will generate continuous oscillations provided that the gain around the loop is greater than one and the proper phase relations are maintained. The output voltage of an *RC*-coupled amplifier having an even number

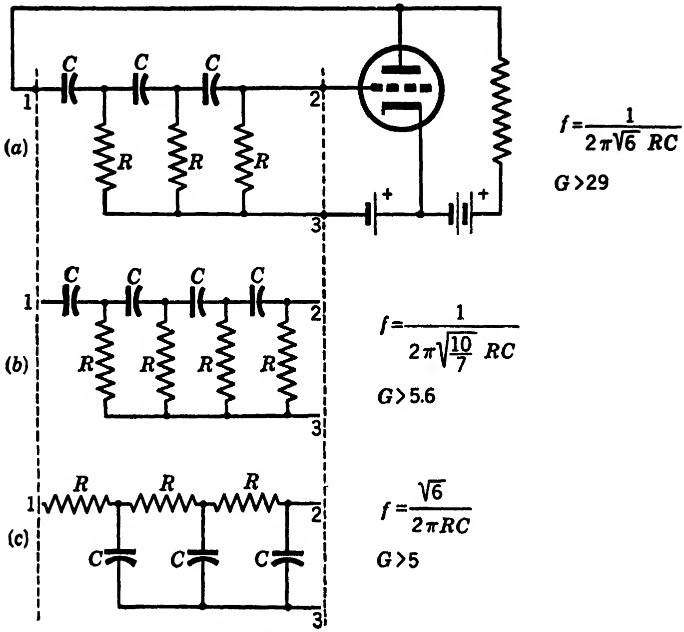


FIG. 7. *RC* phase-shift circuits for oscillators.

of stages will be in the same phase as the input. Hence, if a portion of the output is fed directly back to the input, oscillation will take place. The frequency of this oscillation will be that at which the gain of the circuit is the maximum. Some means must be provided to control the amplitude of oscillations so that the tubes will not operate beyond the linear portion of their characteristics. If this is not done, the amplitude will increase to the point where grid current flows and the output will be greatly distorted. (See multivibrators, page 216.)

A single amplifier stage may be made to oscillate by providing an *RC* coupling having 180 degrees of phase shift between the plate and the grid. When a resistor and capacitor are connected in series across a voltage, the resulting current and potential drop across the

resistor will lead the impressed voltage. Two such networks could provide the maximum shift of 180 degrees but the attenuation would be infinite. A practical circuit, Fig. 7*a*, uses three sections, each with a phase shift of approximately 60 degrees and an over-all attenuation of 29 decibels.* Such values are quite satisfactory because it is not difficult to select an amplifier tube that will provide this gain. More *RC* sections may be used if desired, resulting in a reduction of the minimum required gain as shown in part *b*. The network may also be arranged to have the resistances in series and the capacities in shunt part *c*. This arrangement requires less gain but cannot be used as

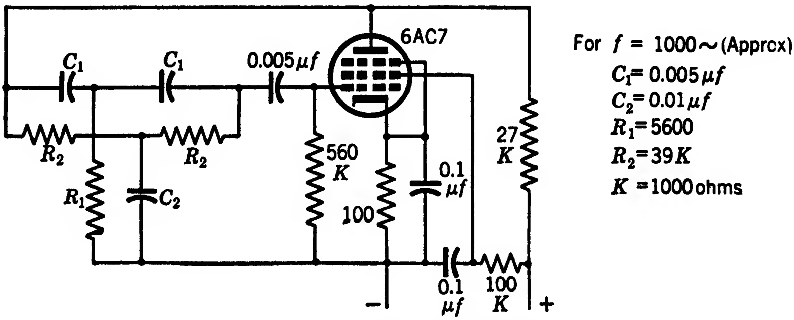


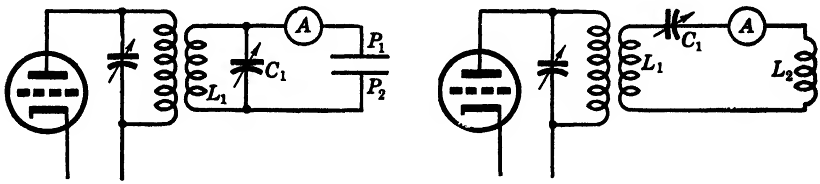
FIG. 8. Parallel T *RC* oscillator circuit.

conveniently since it does not provide a return circuit for the electrons striking the grid. The frequency of these oscillators depends on the *RC* ratio and may be shifted by varying either the resistance or capacity. If a wide frequency range is to be covered, two or more of the circuit elements may be made variable and ganged together. Since frequency varies inversely with R and C ($f = K/RC$), a range of ten to one can be easily covered. In contrast with this flexibility, the resonant frequency in an LC tuned circuit is $f = K/\sqrt{LC}$ and a variation of ten to one in the values of L and C will produce a net change in frequency of the square root of ten or three approximately. The top limit of frequency is governed by the effective interelectrode capacity of the tube but may be as high as 500 kc.† Distortion in the *RC* feedback circuit is low, being 0.1 per cent in a typical circuit. Other circuits, such as the parallel T of Fig. 8, have the required characteristics and make satisfactory *RC* oscillators.

* "Phase-Shift Oscillators," E. L. Ginzton, and L. M. Hollingsworth, *Proceedings I.R.E.*, Vol. 29, p. 43, February 1941.

† "Extending the Frequency Range of the Phase-Shift Oscillator," Rodney W. Johnson, *Proceedings I.R.E.*, Vol. 33, p. 597, September 1945.

Power Oscillators. The power output required from vacuum-tube oscillators depends upon the application. For the majority of applications the oscillator is used merely as a source of a high-frequency signal which is fed into an amplifier where the required power is developed. For such uses a power output from the oscillator of a fraction of 1 watt to 5 watts is ample. This low power output is desirable because larger loads tend to affect the stability and frequency of the Hartley, Colpitts, and similar circuits. The electron-coupled circuit, however, does provide moderate power output with good stability. There are a few applications, such as diathermy and high-frequency heating, where stability of frequency is unimportant but where a large



(*a*) Dielectric (*b*) Induction

Fig. 9. Coupling circuits for rf heating.

power output is needed. For such applications, circuits of the Colpitts, Hartley, or other types may be used with water- or air-cooled tubes and circuit components of high voltage and current capacity so that a power output of 1 to many kilowatts may be attained. Two types of load circuits are used for high-frequency heating as shown in Fig. 9. For dielectric heating, part *a*, the object to be heated is placed between two plates P_1 and P_2 and serves as the dielectric in the condenser formed by these plates. High radio-frequency voltage is applied to the condenser, and heat is generated because of the leakage and other losses in the dielectric. When used for diathermy treatments electrodes P_1 and P_2 are in the form of insulated pads containing a conductive plate or screen. These pads are placed on either side of the portion of the body to be heated. Inductive heating makes use of a coil L_2 which may consist of several turns of an insulated conductor wrapped around the object to be heated. High current rather than high voltage is required here. A coil L_1 couples the load to the tank circuit of the oscillator. The load circuit is tuned by the variable capacitor C_1 in order that maximum power transfer may be obtained.

If power greater than that obtainable from a single tube is required, two or more tubes may be used connected either in a push-pull or in a

parallel circuit. Push-pull circuits are preferred, especially if the operating frequency is high, because the effective capacitances of the tubes are in series across the tank rather than in parallel and thus permit a more favorable ratio of C to L. Tubes for this type of service may be obtained in sizes ranging from those having a rated dissipation of only a few watts to those with water-cooled plates which can dissipate 100 kilowatts or more. A typical tube used in diathermy machines is pictured on page 71. This tube has a rated plate dissipation of 125 watts and is normally operated with a plate voltage of 1500 volts and a current of 200 milliamperes. Two of these tubes in a push-pull or parallel circuit will deliver over 400 watts of power at frequencies up to 30 megacycles. A much larger tube for use in radio-frequency heating machines is pictured on page 75. This tube is designed for forced-air cooling and is similar to water-cooled tubes except for the heavy copper fins which cool the plate. Normal plate voltage is 18,000 and with a current of 3.6 amperes the output will be about 50 kilowatts. Power to heat the filaments is usually obtained from a step-down transformer. The smaller tube requires only 3.25 amperes at 10 volts, whereas the larger one employs 61 amperes at 10 volts. Plate voltage is usually supplied from a rectifier system (see Chapter XIV) although sometimes the rectifier is omitted and alternating voltage is used directly from a step-up transformer. Such operation is not recommended because plate current can flow only when the supply voltage is positive and because such operation causes a large amount of hum modulation. Efficiency and power output are consequently lower, and the spurious signals generated may interfere with other services.

In the operation of oscillators, it may happen that a part of the circuit will oscillate at one or more frequencies in addition to the desired frequency. Such spurious frequencies are due to *parasitic oscillations*. Parasitic oscillations arise from resonant conditions which may exist in the components of the amplifier and oscillator circuits. These conditions are generally due to the combinations of interelectrode capacities with other components of the circuit. Two or more power tubes in parallel increase the likelihood of parasitic oscillations because more components and more interelectrode capacities are involved. Parasitic oscillations are less likely to occur in screen grid and pentodes because of the more effective isolation of grid, cathode, and plate.

Oscillations in a power oscillator may be blocked under certain unfavorable conditions. Blocking is caused by secondary emission from the grid which results in a reversal of electron flow and a reversal

or polarity in the grid-bias circuit. Thus the grid is driven positively, and a large plate current will result which will soon destroy the tube. Blocking may be caused by a combination of high plate-load resistance, high grid-leak resistance, and excessive feedback signal. Most modern tubes are so constructed and the grids so treated that secondary emission is restrained and blocking is unlikely except in very high voltage operation.

Negative Resistance Oscillators. A negative resistance oscillator is one in which oscillations are sustained through negative resistance in connection with an *LC* tank circuit. Oscillations once started in a

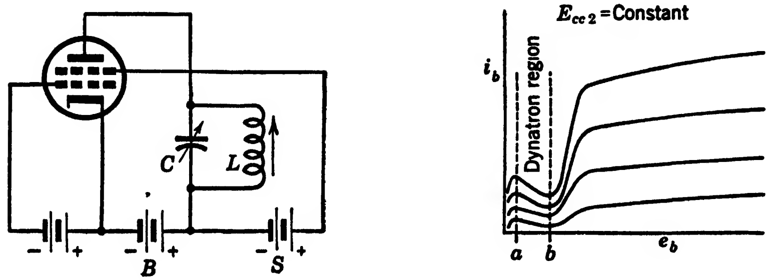


FIG. 10. Dynatron oscillator circuit and characteristic.

resonant *LC* circuit containing zero resistance would continue indefinitely. Although zero resistance cannot be realized within the *LC* closed circuit, it is feasible to associate a negative resistance external to the *LC* circuit. Negative resistance characteristics are found in the operation of multielectrode tubes such as the tetrode and pentode.

A negative resistance oscillator circuit employing a screen-grid tube and known as a dynatron oscillator is shown in Fig. 10. The plate current-plate voltage curves of this figure show the negative resistance (dynatron) region in which the circuit must function. In order to sustain oscillations the negative resistance ($\Delta e_b/\Delta i_b$) should be greater than the resistive component of parallel impedance of the *LC* tank circuit at the resonant frequency.

A physical concept of the action of the dynatron may be obtained from Fig. 10. The number of primary electrons reaching the plate of the tube in the "dynatron region" is independent of the plate voltage, but the number of secondary electrons emitted from the plate increases with the plate voltage. If the screen is more positive than the plate nearly all the secondary electrons will flow to the screen, thence through a portion of the battery *S*, *L* of tank, and back to plate. The potential from cathode to plate e_p is the battery *B* plus the drop across

the LC tuned circuit. As e_p swings from position a to b on the curve the electrons move through L in the direction shown by the arrow and they are decreasing in number, giving a negative resistance characteristic. The electron current is produced by battery S, and it serves to store energy in the capacitor of the LC circuit. When e_p tends to move to the right of position b owing to energy stored in L, the change to a positive resistance brings the current to a halt and the energy stored in C swings e_p back to position a where the resistance again becomes positive and arrests the swing. For these oscillations to be sustained, it is necessary for the negative resistance of the tube in the dynatron region to be greater than the positive resistance of the LC tuned circuit.

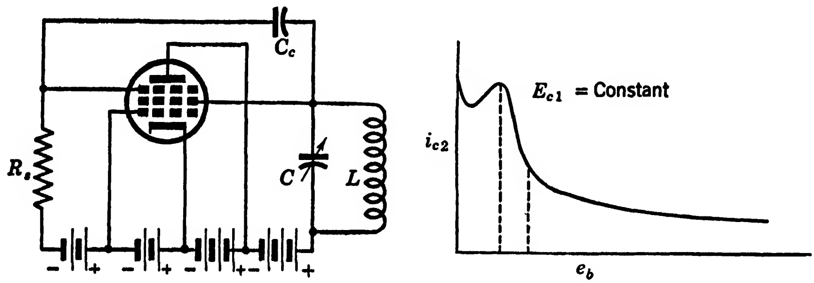


FIG. 11. Transitron circuit.

Tubes in which the plate has received special treatment to prevent secondary emission will not serve for a dynatron oscillator, but a tube such as that shown on page 80 will serve satisfactorily. Oscillators operating on the dynatron principle are not well suited for large power output but have very stable frequency characteristics and make good frequency standards.

Negative resistance also occurs in the screen-grid circuit of a pentode when its suppressor and screen grid are coupled together. If an LC tank circuit is connected to the screen as shown in Fig. 11, oscillation will result. This circuit is called the transitron or negative-transconductance oscillator. It is more reliable than the dynatron because it does not depend on secondary emission but functions through the negative resistance of the i_{c2}-e_{c2} characteristic.

Mechanical Resonance for Oscillators. The operating frequency of an oscillator may be controlled by mechanical resonance instead of the resonance of an electrical circuit. One system for mechanical control of the exciting frequency is illustrated in Fig. 12. Here the frequency is established by the natural period of vibration of a tuning

fork. The conversion of the mechanical resonance to an electrical signal is performed in the microphone, transformer, and battery circuit. Losses are supplied and vibration is maintained by the periodic impulses given by the electromagnet M. The precision of the frequency is determined by the characteristics of the tuning fork and the ambient temperature.

At audio frequencies a tuning fork with an electrical pickup and driver will hold its frequency to within 0.05 per cent. Frequency multiplying circuits may be used to provide multiples of the base frequency of the tuning fork. This system of frequency generation has been employed in carrier telephone circuits.

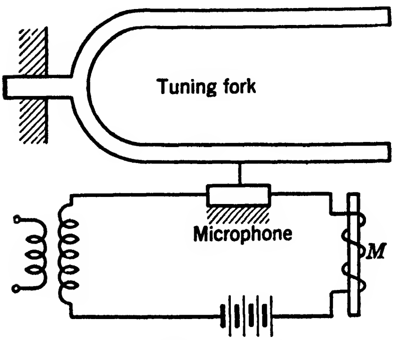


FIG. 12. Tuning-fork frequency generation.

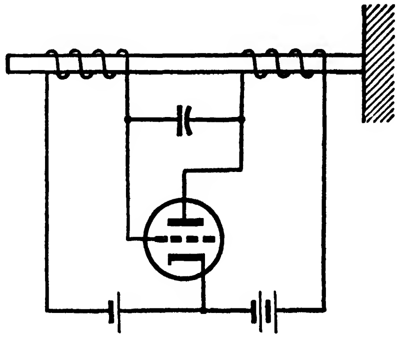


FIG. 13. Magnetostriction oscillator circuit.

Another mechanical system of frequency generation employs the phenomenon of magnetostriction. Changes in magnetism in some metals produces an expansion or contraction of the metal. This property is utilized in the circuit of Fig. 13 where the coil on one end of a rod of magnetic material is connected in the plate circuit of a triode while a coil on the opposite end feeds the grid circuit. A change of current in the plate circuit produces a contraction or expansion of the rod at the right which sends a compression wave to the left. As this wave is propagated to the left it varies the magnetism in the rod and induces a voltage into the grid circuit. The resulting signal on the grid will produce a change in the plate circuit. If the electrical circuits are tuned to the period of travel of the wave owing to magnetostriction and if phase relations are correct, the circuit will oscillate. It is apparent that this oscillator represents a form of feedback, but it is not an inductive feedback (transformer action), a capacitive feedback, or a direct circuit feedback. The feedback arises from the

magnetostrictive property of metal. These oscillators operate at audio and supersonic frequencies and are constant to approximately 0.03 per cent.

In passing it should be noted that frequency generation in the crystal oscillator utilizes the period of mechanical resonance of the crystal though the piezoelectric property is a combination of mechanical and electrical properties.

Nonsinusoidal Oscillators. For many applications wave forms are desired that are not sinusoidal. Such waves are usually rectangular or triangular with respect to the time axis. They are used for control

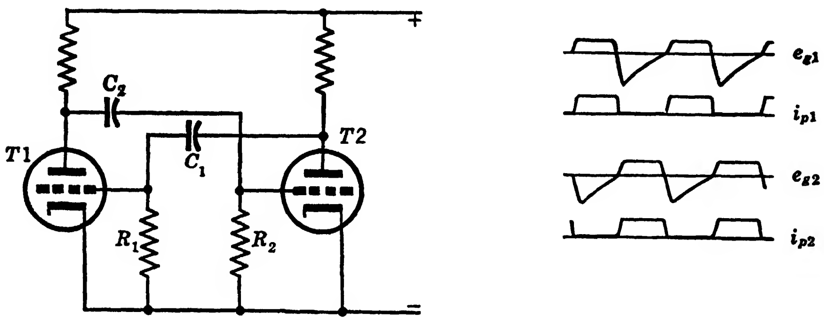


FIG. 14. Multivibrator circuit.

purposes where it is necessary to have a circuit turn on and off periodically, for the time base in oscilloscopes, for the sweep circuits in television receivers, for pulse and marker circuits in radar systems, and for many other applications. In the circuits that generate oscillations of this type, the tube operates as an electronic switch which opens and closes to control the current flow through a load resistance. Usually the grid is driven far into the positive region so that the tube operates under Class C conditions, with the operating angle adjusted to give the required wave shape.

The most common nonsinusoidal oscillator is the multivibrator which consists of a two-stage, RC-coupled amplifier whose output is fed directly back to the input, as illustrated in Fig. 14. The frequency of oscillation is determined by the time constant of the coupling condensers and grid resistors, and, if the time constants in the two grid circuits are equal, the wave will be symmetrical. An increase in voltage at plate $T1$ causes grid $T2$ to swing toward the positive, which increases the current in $T2$. This rise of current lowers the plate voltage of tube $T2$ and the grid voltage of $T1$. The reduced grid potential of $T1$ causes the voltage at plate of tube $T1$ to rise even higher, and

the action proceeds until grid of $T2$ is driven positively and grid of $T1$ negatively past cutoff. Grid of tube $T1$ will lose its negative charge at a rate determined by its RC circuit. As soon as tube $T1$ begins to draw plate current the action reverses very rapidly until grid of $T1$ is driven positively and $T2$ is negative. The rate at which oscillation occurs is approximately $1/(R_1C_1 + R_2C_2)$ cycles per second. Plate current flows in rectangular pulses, first in one tube and then in the other. The multivibrator circuit may be adjusted for frequencies from less than 1 to 100,000 cycles per second.

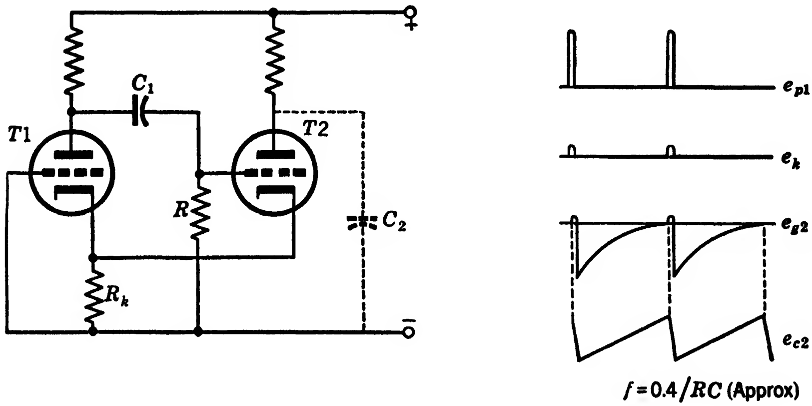


FIG. 15. Potter oscillator circuit.

Multivibrators are useful because their output contains harmonics up to several hundred times the fundamental frequency. Also, they are sensitive to synchronization since their rate of oscillation may be readily increased by the application of a small triggering voltage. Under these conditions the multivibrator frequency will be the synchronizing frequency divided by an integer. When the integer is greater than unity, the circuit acts as a frequency divider. Stability decreases as the rate of division increases but is satisfactory up to ratios of ten or more.

A circuit invented by J. L. Potter * uses a common cathode impedance R_k in place of one of the grid-plate coupling condensers, Fig. 15. Plate current in tube $T2$ is in the form of a sharp pulse and may be used to discharge a condenser C_2 and provide a sawtooth voltage, e_{c2}. Synchronization can be readily applied to grid of $T1$ since no other connections are necessary to this point. Another circuit for produc-

* J. L. Potter, "Sweep Circuit," *Proceedings I.R.E.*, Vol. 26, p. 713, June 1938.

ing sharp current pulses is the blocking oscillator shown in Fig. 16, which uses an iron-cored transformer connected for positive feedback. Increase in the plate current of the tube induces a voltage in the grid winding of the transformer which drives the grid far into the positive region; thus a large grid current is caused to flow and charge the coupling condenser. Plate current increases until the tube reaches saturation. At this point the large negative grid voltage which has been developed across C_g begins to fall as C_g discharges through R_g. Plate current cannot flow again until the charge on the capacitor leaks

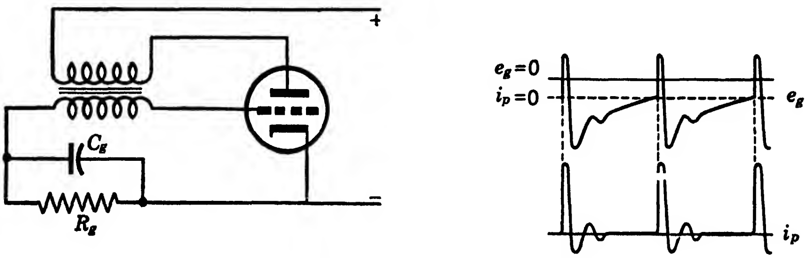


FIG. 16. Blocking oscillator circuit.

off through the grid resistor. The duration of the plate-current pulses depends mainly upon the natural resonance of the feedback transformer, whereas the rate at which these pulses occur is determined by the RC constant of the grid circuit. Blocking oscillators can be readily synchronized and may be used as frequency dividers. They may be followed by a discharge tube and condenser to provide a sawtooth wave form. Both the Potter and the blocking circuits are well suited to serve as the sweep oscillators in television receivers or to provide the time base for oscilloscopes. Blocking oscillators can also be used to advantage in radar equipment to establish the pulse repetition rate and to generate marker "pips."

Oscillators whose frequency is controlled by the charge or discharge of a condenser through a resistance are sometimes called "relaxation oscillators." One simple relaxation oscillator circuit is shown on page 108. There are several other types of relaxation circuits which use screen-grid tubes or gas tubes (thyratrons). The latter is widely used in oscilloscopes and will be discussed in a later chapter.

Chapter X

MODULATION AND DETECTION

Modulation and Carrier Currents. In the general sense the term modulation means to mold or form an electric current to conform to some variable. This concept is illustrated in Fig. 1 which shows one simple circuit (*a*) in which the dots and dashes of the Morse code are translated into rectangular pulses of current and voltage. In a similar manner, sound waves impressed on a carbon granule type of micro-

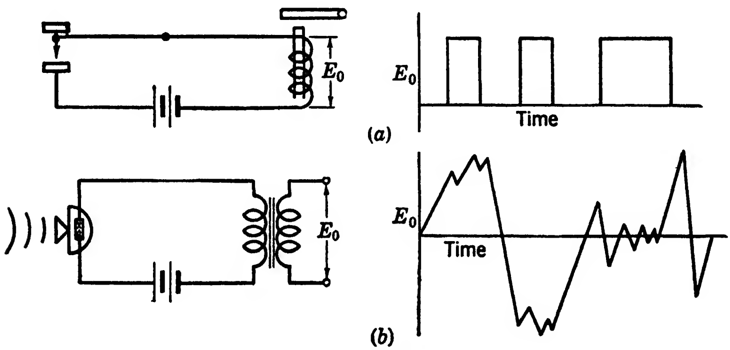


FIG. 1. Modulation in simple telegraph and telephone circuits.

phone (telephone transmitter) will vary its resistance and produce a voltage output in the transformer of Fig. 1*b* which has a voltage amplitude and frequency characteristic that conforms to the sound signal. Suitable detecting or receiving instruments such as telegraph sounders or telephone receivers may be employed to translate the modulated voltage and current waves back into replicas of the original motion or sound.

In a more specific sense modulation is defined as the process of producing a wave some characteristic of which varies as a function of the instantaneous value of another wave.* As applied in electronic circuits modulation is employed in connection with high-frequency currents known as carriers. The term carrier follows from the fact

* I.R.E. Standards on Transmitters and Antennas, 1938.

that intelligence is transmitted by impressing signal currents upon a high-frequency current by some process of modulation. High frequencies are employed for carriers for two reasons. First alternating currents cannot be propagated satisfactorily or efficiently via radiated waves (radio) at lower frequencies. Thus a 60-cycle current may require an antenna 60 miles long to secure much radiation. At frequencies above 15,000 cycles per second, alternating currents may be propagated through space more satisfactorily. The second reason for the use of high frequencies for carriers is the fact that a low-frequency signal may be impressed on a carrier (modulation) in a way to require a rather narrow percentage band width of frequencies. Thus the utilization of different frequency bands provides multiple channels for transmission through space (radio) or several channels for telephone messages over a single pair of wires.

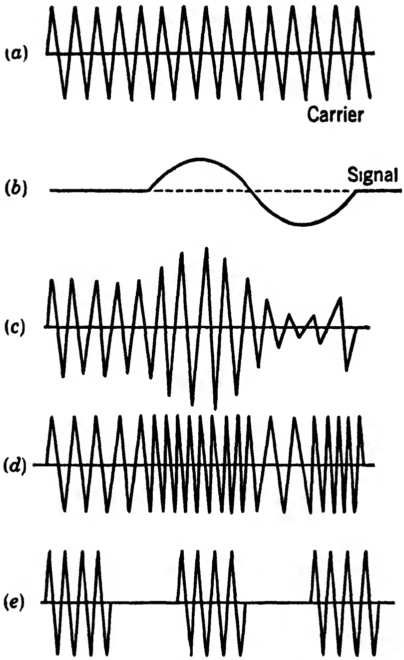


FIG. 2. Carrier signal and modulated waves.

The application of modulation to carrier currents is illustrated in Fig. 2. When a low-frequency signal such as b of Fig. 2 is impressed in some suitable manner on the high-frequency carrier of a, a modulated wave as suggested in parts c or d results. The simplest type of modulation of a carrier is produced by an opening or closing of the carrier circuit in conformation with the dots and dashes of the Morse code, as illustrated in part e of Fig. 2. Other types of modulation based on the information being transmitted are radio telephone, facsimile, and television. Radio telephony includes both speech and music. Facsimile is the transmission of still pictures, and television is the transmission of pictures at a rate sufficient to give the illusion of motion.

Classes of Modulation. A sine wave of carrier current can be represented by the equation

$$i = A \sin (2\pi ft + \phi) \tag{1}$$

where A represents the maximum value or amplitude, f the frequency of the periodic wave, and ϕ the phase angle with respect to some fixed reference. The form of the wave represented by equation 1 can be changed by introducing variations with time in the values of either A, f, or ϕ. Such variations give rise to three of the important forms or classes of modulation of carrier currents which are termed *amplitude* AM, *frequency* FM, *phase* PM modulation. In each class the factor causing the variation is the a-c signal wave which is to be transmitted. Amplitude modulation is illustrated in parts c and e of Fig. 2, and frequency modulation and phase modulation is pictured in part d.

One concept of the three classes of modulation may be obtained by reference to the operation of an a-c generator. First, if the rotating member of an alternator is driven at constant speed and the field current is varied with time, the amplitude of output voltage of the generator will follow the field changes but the frequency will remain constant. The resulting wave form corresponds to amplitude modulation. Second, if the speed of the alternator is varied with time and if simultaneously the field of the alternator is varied also in a manner to maintain a constant maximum voltage, then the resulting frequency of the output must vary though the amplitude of the waves is constant. This analogy represents frequency modulation. Third, if the speed and the field of an alternator is kept constant but if the stationary part of the alternator is movable so that it can be oscillated back and forth through a small angle, then the factor ϕ in equation 1 may be changed. The resulting wave form of the alternator represents phase modulation.

A further concept of modulation may be obtained from Fig. 3. In part a, if the carrier vector I_c rotates at a uniform rate about point 0, its projection on the vertical will generate the carrier wave shown at the right with the instantaneous value of the carrier current

$$i_c = I_c \sin \omega_1 t \tag{2}$$

Now if a modulating signal current of value $i_m = I_m \sin \omega_2 t$ is impressed upon the rotating vector of part a, it will result in the action and wave depicted in part b. Thus the length of the rotating vector will oscillate between the extreme positions of a and b, and the projection on the vertical will generate the amplitude modulation wave shown on the right. The instantaneous value of points in the modulation wave is

$$i = (I_c + I_m \sin \omega_2 t) \sin \omega_1 t \tag{3}$$

It will be noted that the modulating signal causes amplitude changes at right angles to the time axis. By way of contrast, frequency and

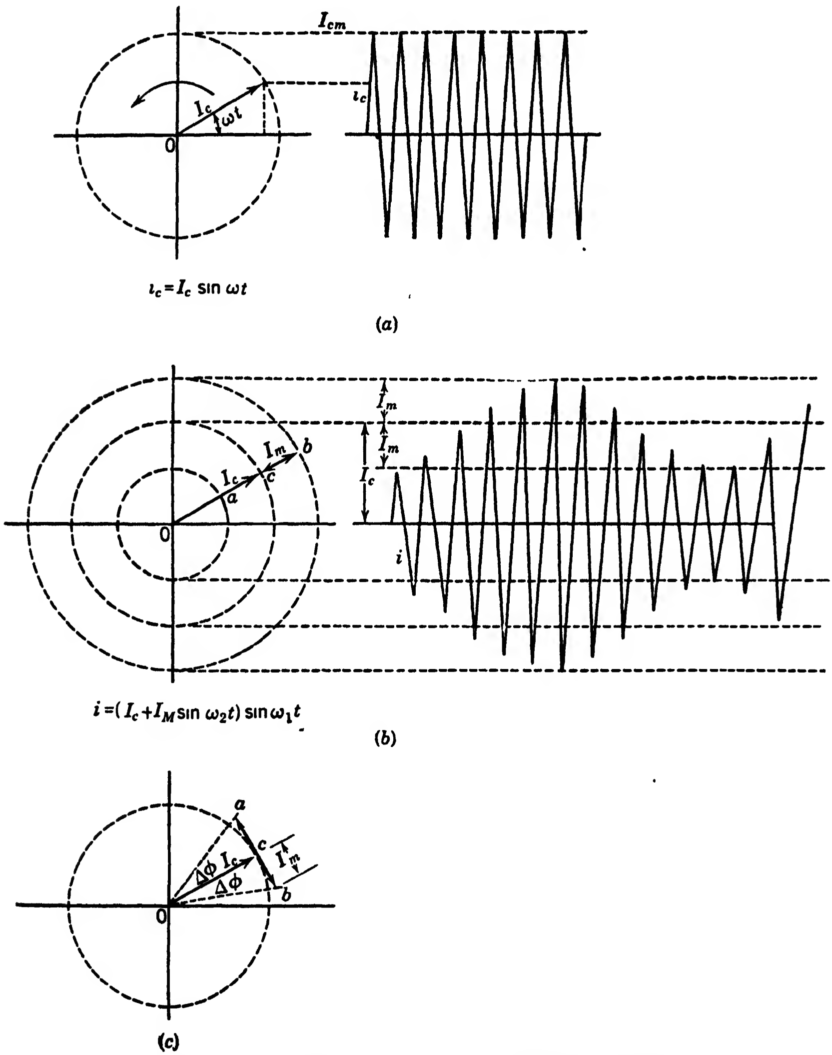


FIG. 3. Generation of carrier and modulated waves.

phase modulation produce changes "along" the time axis. Thus amplitude and frequency modulation act in quadrature in respect to the time axis.

A further concept of phase modulation is given in Fig. 3*c*. The carrier vector I_c is caused to oscillate back and forth through some vary-

ing angle $\Delta\phi$ while it is in the process of rotation about point 0. The generating vector projected on the vertical and a time axis will look similar to the representation of frequency modulation.

Nature of Modulation Signals. The term modulation factor for amplitude modulation shown in Fig. 3*b* is defined by the equation

$$m = \frac{I_{\max} - I_{\min}}{2I_c} \tag{4}$$

This equation represents the maximum change in amplitude divided by two times the average amplitude. In general, this is the ratio of the maximum value of the modulating signal to the maximum value of the carrier. The ratio m times 100 is called the per cent of modulation.

In amplitude modulation the sinusoidal carrier current may be represented by the expression

$$i_c = I_c \sin \omega_1 t \tag{5}$$

If the carrier is modulated by a signal of value m,

$$i_m = mI_c \sin \omega_2 t \tag{6}$$

the resulting modulated current will be

$$i = I_c(1 + m \sin \omega_2 t) \sin \omega_1 t \tag{7}$$

The expansion of this expression and substitution of equivalent trigonometric terms for the $\sin \omega_1 t \sin \omega_2 t$ gives for the modulated current

$$i = I_c \sin \omega_1 t + \frac{m}{2} I_c \cos (\omega_1 - \omega_2)t - \frac{m}{2} I_c \cos (\omega_1 + \omega_2)t \tag{8}$$

This expression shows that the amplitude-modulated carrier is made up of three components. One component is the original carrier, and the other two arise from the sum and difference frequencies of the carrier and signal. For example, if the carrier is 100,000 cycles and the signal 10,000 cycles, the modulated carrier will contain components of the carrier—100,000 cycles, 100,000 + 10,000 = 110,000 cycles, and 100,000 − 10,000 = 90,000 cycles. *The sum and difference frequency components are called the side bands.* There are two side-band components for each frequency in the signal. Thus, if the signal contains a range of frequencies from 40 to 5000 (voice transmission), a large range of component sum and difference frequencies will exist in the

side bands. It is obvious that the intelligence must be transmitted by the side bands because the carrier component does not contain any imprint of the signal.

For 100 per cent modulation ($m = 1$), it follows from equation 8 that the side-band frequencies will have an amplitude one-half that of the carrier. If the carrier power in the output is considered as 100 per cent, then the output power for each side band is $(I_c/2)^2R$, or 25 per cent of the carrier output where R is the resistance of the load. Thus the total output of the modulated carrier is 150 per cent of carrier power and the carrier contributes two-thirds and each side band

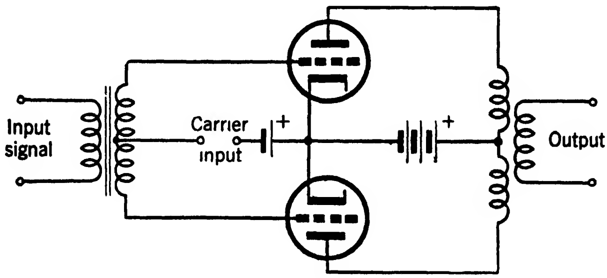


Fig. 4. Balanced modulator circuit to suppress carrier.

one-sixth of the total output. Since the intelligence is contained in either side band, it is evident that a saving in power could be effected by suppressing the carrier and one side band. A method for suppressing the carrier is suggested in the simple balanced modulator circuit of Fig. 4. This is a push-pull circuit wherein the carrier input will be canceled in the primary of the output transformer. One side band can be removed by a filter. The suppression of the carrier and sometimes one side band is used in transoceanic telephony and long distance telephone circuits. However, the carrier is not suppressed in radio broadcast systems because the necessary receiving equipment for this method of transmission requires a synchronized source of carrier supply and hence is costly and complicated.

With amplitude modulation the minimum value of the modulated current can never be less than zero; and the maximum value I_{max} cannot exceed twice the unmodulated carrier value I_c if a sinusoidal relationship is maintained. However, with complex modulating signals having positive peaks higher than the negative, I_{max} may exceed $2I_c$ without introducing distortion and the modulation factor must then be redefined as follows: for positive modulation peaks

$$m = \frac{I_{\max} - I_c}{I_c} \tag{9}$$

and for negative modulation peaks

$$m = \frac{I_c - I_{\min}}{I_c} \tag{10}$$

The general form of the equation for the current modulated with frequency modulation is

$$i = I_c \sin (\omega_1 t + B \sin \omega_2 t) \tag{11}$$

The solution of this equation involves Bessel functions and is beyond the scope of this textbook. It should be observed that the main sine function consists of the carrier variation $\omega_1 t$ plus a component which

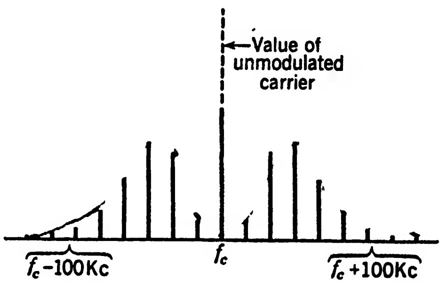


FIG. 5. Side band components of a frequency modulation wave.

varies as the sine of the signal. The resultant change produces the deviation of frequency which constitutes frequency modulation. The resultant current consists of a carrier component plus side-band frequencies. However, these frequencies bear an entirely different relationship to the modulation signal and the carrier.

In amplitude modulation there are only two side bands, whereas in frequency modulation there are several, the number depending on the swing or variation of the carrier frequency from the normal value. A concept of the carrier and side bands is given in the frequency spectrum shown in Fig. 5. The large number of side bands in FM means that a relatively wide band is required in the frequency spectrum for an FM channel. Since the amplitude of the modulated carrier remains constant, there is no change in power output during modulation. The side-band energy must be subtracted from the carrier. With amplitude modulation the per cent of modulation was defined as the per cent of deviation in carrier amplitude. In FM the per cent of modulation

is based on some arbitrary value of frequency deviation. For FM broadcasting service this deviation or variation of the carrier from its mean value has been set at ±75 kc. Per cent of frequency modulation is the actual frequency deviation divided by 75 kc times 100. The maximum deviation in the negative direction would swing the carrier from its normal value to zero. Similarly, a positive deviation might be enough to push the carrier frequency to several times its normal value. Practical circuit limitations, however, dictate a deviation of only a few per cent of the carrier if severe distortion due to nonlinearity in the rf circuits is to be avoided.

In amplitude modulation, the amplitude of the modulation signal is carried by the amplitude of the modulated carrier; the pitch or frequency of the signal is conveyed by the number of changes in amplitude per second. In FM, the amplitude of the modulation signal is carried by the per cent of frequency modulation; the pitch is conveyed by the number of frequency changes per second similarly to AM. In PM the amplitude of the modulation signal is carried by the degree of phase-angle change imposed upon the carrier; the pitch is conveyed as in AM and FM. In both FM and PM the varying factor is a sinusoidal function and the resulting modulated wave suffers both a frequency and a phase-angle change. In practice, it is impossible to produce one without the other being present. Thus the distinction between FM and PM is more a matter of definition of terms than one of practical significance. However, there is a real difference as far as the side bands are concerned.

Modulation Methods and Circuits. Amplitude modulation may be produced by many different methods and circuits. The simplest and oldest method is by absorption. A variable resistor such as a carbon granule microphone may be placed in series with the antenna circuit as illustrated in Fig. 6, where its varying resistance will control or modulate the output carrier power. Absorption modulation was widely used in the early days of radio telephony but is without application today. The principle of modulation by rectification is illustrated in Fig. 7. Here the carrier and the modulation signal are fed in series into a diode with a tuned-circuit load impedance. With zero modulation signal the diode will rectify the positive loops of carrier

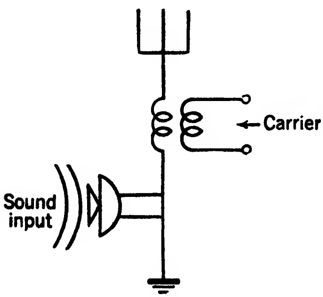


FIG. 6. Simple circuit for absorption modulation.

current and all will have the same amplitude. Added modulation signal variations will change the magnitude of the resultant cathode-plate voltage and produce plate-current rectified loops of varying amplitude as shown on the right of Fig. 7. The current flowing in the tuned-plate circuit will have full waves with the negative loops supplied by the tuned-tank circuit. A transformer coupling applied to the tank circuit would supply a complete amplitude modulated wave for application to an amplifier.

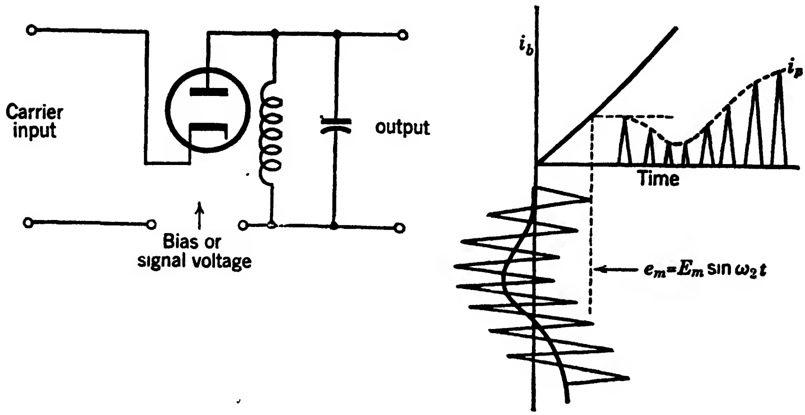


FIG. 7. Amplitude modulation using a diode.

A crystal rectifier or diode (see Chapter XIII) can be substituted on high-frequency circuits in place of the vacuum diode of Fig. 7. This principle of modulation by rectification has the disadvantage that the modulated carrier is of low magnitude because amplification is not involved in the circuit employed. Accordingly, this principle finds little application in radio circuits but is employed in a modified form in telephone circuits where currents of small magnitudes are the rule.

The most widely used circuit for amplitude modulation employs plate modulation, as illustrated in Fig. 8. Here the carrier is fed into the grid of a Class C amplifier while the modulation signal is fed into the plate circuit. The modulation signal in series with the plate-voltage supply causes the resultant plate voltage to be $E_{bb} \pm E_m \sin \omega_2 t$, where E_m is the maximum value of the signal voltage. The result of this plate-voltage variation is to cause the amplification of the circuit to vary along the i_b-e_b dynamic characteristic of the tube as suggested in the right-hand view in Fig. 8. Since the dynamic characteristic is

nearly linear over a fairly wide range of operation, the output will be nearly linear and with a low degree of distortion.

Plate modulation is possible because a Class C amplifier or oscillator acts like a pure resistance as far as the B+ supply is concerned. If the plate voltage is doubled, the current will be doubled likewise and the d-c power input to the stage will increase in the ratio of 4 to 1. If the circuit is properly adjusted, efficiency will remain constant and the power output will vary directly as the d-c input. Since the average output for 100 per cent modulation must be one and one-half times the carrier value, it is necessary that the device causing modulation

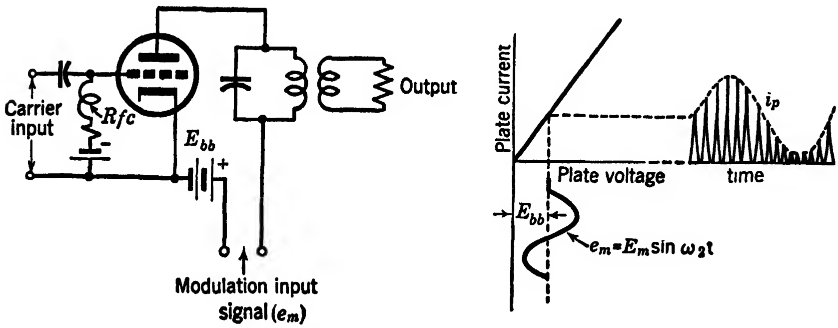


FIG. 8. Plate amplitude modulation using Class C amplifier.

be able to supply 50 per cent as much power as the normal input to the Class C stage plus some power for losses in the modulating tube. This device is called the *modulator* since its function is to vary the plate voltage of the rf stage and thereby control the output. The modulator consists of a Class A or AB_1 or AB_2 or push-pull Class B amplifier coupled to the Class C rf stage, as shown in Fig. 8. The modulator must be capable of passing the frequencies which are to be impressed upon the carrier. In a broadcast transmitter handling speech and music, the modulator must pass frequencies from about 40 cycles per second to 15,000. In large transmitters where the carrier power is 50 kilowatts to the final rf amplifier, the modulator must be able to deliver 25 kilowatts of audio power.

Amplitude modulation may originate in the grid circuit of an amplifier operated as Class A, B, or C. Class C operation is shown in Fig. 9, where the modulation signal is impressed on the grid in series with the carrier. The modulating voltage serves to vary the grid bias and thus control the amplitude of the output plate current as shown. The current flow in the tuned-plate circuit will be full wave as depicted in the lower right view of the figure.

For grid modulation, the stage must be adjusted so that its output current varies linearly with change is bias. The amplifier acts like an underexcited Class C stage under zero modulation and in true Class B fashion on modulation peaks. E_{bb} remains constant, but when $m = 1$ both the plate current and the efficiency double; thus a peak power output of four times the carrier value is supplied.

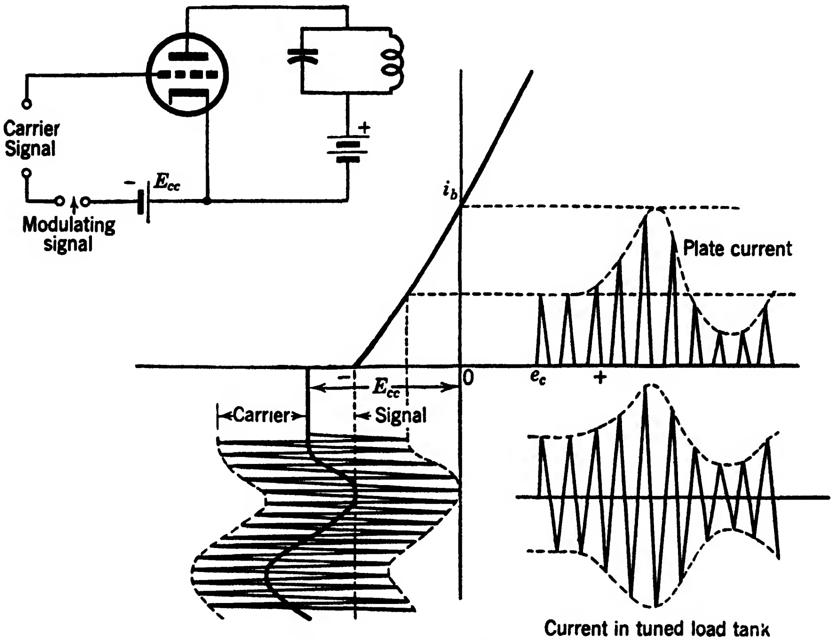


FIG. 9. Grid-modulated Class C amplifier.

Grid modulation requires much lower audio power than plate modulation does with a resulting saving in size and cost of circuit components. Because of the low power requirements, it is practicable to use resistance coupling whenever high modulating frequencies are encountered, as in television transmitters. These advantages are offset by the lower efficiency (about 33 per cent) of the grid-modulated stage and by the added requirements placed on the rf driver. Over the modulation cycle the rf load presented by the grid varies, and, in order to maintain constant rf grid voltage, the driver must have low impedance. Usually a resistive load is connected across the grid circuit to consume most of the driver power. Circuit adjustment is somewhat difficult since bias, audio grid voltage, rf drive, and plate load must all be right if low distortion and reasonable power output are to be obtained.

If a pentode is used as the rf amplifier, modulation may be accomplished by applying the audio signal to the suppressor and the rf drive to the control grid. Operation is essentially the same as for a grid-modulated triode.

Class B rf stages, when properly adjusted, may be used as "linear" amplifiers. Output current varies directly and linearly with rf input

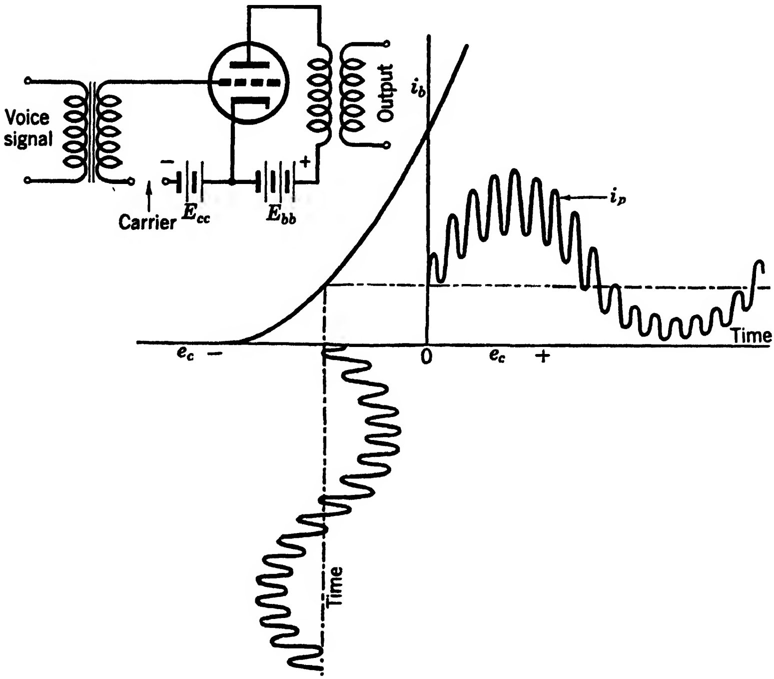


FIG. 10. Grid modulation of Van der Bijl type.

voltage so that a modulated signal impressed on the grid will be amplified without distortion. Efficiency is about 33 per cent, and the driver must have low impedance as with grid modulation. One or more linear amplifiers may be used following a low-power plate or grid-modulated stage to provide greater output.

A transmitter which is grid-modulated or which uses linear output stages is said to be "low-level modulated" since modulation occurs at a point where the rf power is relatively small. On the other hand, if modulation is at the plate of a Class C final amplifier, it is "high-level." Because of the better over-all efficiency, inherently low distortion, and ease of adjustment, high-level modulation is usually prefer-

able. The most powerful broadcast transmitters, both medium and short-wave, use high-level modulation with carrier powers of 500,000 watts or more.

One form of grid modulation known as the Van der Bijl type uses a Class A amplifier circuit in which an rf carrier of low magnitude and an audio signal of larger magnitude are applied to the grid. This form of modulation has had wide application in carrier-current telephone communication. It requires a small amount of carrier and signal power and gives satisfactory operation where a small degree of modulation is necessary. The circuit and operation of this system is illustrated in Fig. 10. The nonlinear characteristic of the plate-current curve produces several components of modulation in the resulting output current. This output current i_b may be analyzed by the assumption (approximately true) that the grid-voltage plate-current curve is parabolic in form. Thus

$$i_b = K(E_b + \mu E_c + \mu e_g)^2 \tag{12}$$

$$= K[(E_b + \mu E_c)^2 + 2(E_b + \mu E_c)\mu e_g + \mu^2 e_g^2] \tag{13}$$

Assuming that the voice signal voltage is represented by $A \sin \omega_2 t$ and the carrier voltage by $B \sin \omega_1 t$, the varying grid voltage becomes

$$e_g = A \sin \omega_2 t + B \sin \omega_1 t \tag{14}$$

and

$$e_g^2 = A^2 \sin^2 \omega_2 t + 2AB \sin \omega_2 t \sin \omega_1 t + B^2 \sin^2 \omega_1 t \tag{15}$$

Substituting the trigonometric equivalents for $\sin^2 \alpha$ and $\sin \alpha \sin \beta$, equation 15 may be transformed to

$$e_g^2 = \frac{1}{2}(A^2 + B^2) - \frac{A^2}{2}\cos 2\omega_2 t - \frac{B^2}{2}\cos 2\omega_1 t$$
$$+ AB\cos(\omega_2 t - \omega_1 t) - AB\cos(\omega_2 t + \omega_1 t) \tag{16}$$

A substitution of equations 14 and 16 in equation 13 and a collection of terms reduces to

$$i_b = K(E_b + \mu E_c)^2 + \frac{K\mu^2}{2}(A^2 + B^2)$$
$$+ 2K\mu(E_b + \mu E_c)[A \sin \omega_2 t + B \sin \omega_1 t]$$
$$- \frac{K\mu^2}{2}[A^2 \cos 2\omega_2 t + B^2 \cos 2\omega_1 t]$$
$$+ K\mu^2 AB[\cos(\omega_1 - \omega_2)t - \cos(\omega_1 + \omega_2)t] \tag{17}$$

The analysis of this equation shows the following components of current:

Line 1—direct current.

Line 2—amplified voice and carrier signals.

Line 3—double frequency of voice and carrier.

Line 4—the sum and difference frequencies (side bands).

The direct current of line 1 does not appear on the output side of the transformer in the plate circuit. The magnitude of the voice and carrier components of line 2 may be reduced by giving E_c a large nega-

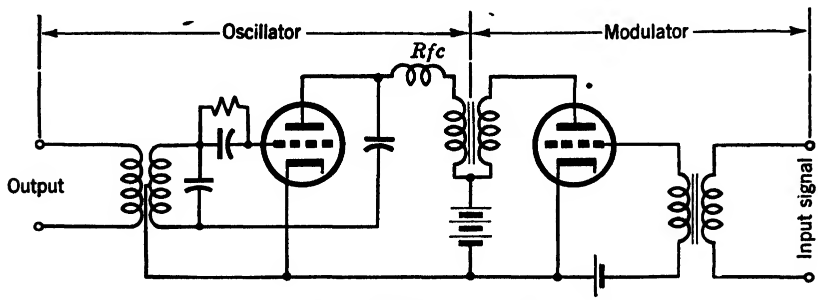


FIG. 11. Plate-modulated oscillator circuit.

tive value. The double frequency components and, if desired, one side band may be removed by a filter.

A later and common form of modulator circuit for telephone carrier systems is illustrated in Fig. 4. This balanced modulator circuit is of the "push-pull" type wherein the carrier is fed into the common input lead to the cathodes. This connection causes the carrier and the double frequencies to be suppressed in the output transformer so that only the signal frequency and the two side-band voltages appear in the output of the modulator.

Amplitude modulation may be produced by introducing the modulation signal in the plate supply circuit for an oscillator tube as shown in Fig. 11. The output voltage of the plate circuit of an oscillator varies as the plate supply voltage; hence this system should give linear modulation. Unfortunately, the frequency produced by an oscillator tends to vary with the plate voltage, and thus this method of modulation gives some frequency as well as amplitude modulation. For this reason the plate-modulated oscillator circuit has limited application.

Frequency modulation can be effected by changing the reactance of the tuned circuit which controls the frequency of a self-excited oscillator. Since the oscillator operates at the resonant frequency of its

LC tuned circuit, any device placed in parallel with the *LC* branch that varies the resultant value of either *L* or *C* will vary the frequency of the oscillations. The simplest device for frequency modulation by an audio signal is a condenser microphone. A condenser microphone consists of a thin sheet of light metal (Duralumin) stretched tightly a short distance from a parallel insulated metal plate. Sound waves falling on the diaphragm cause it to vibrate and the resulting motion varies the capacitance between the diaphragm and the plate. If the diaphragm and plate are coupled to the *LC* tuned circuit of an oscillator, frequency modulation will result. A second method of varying

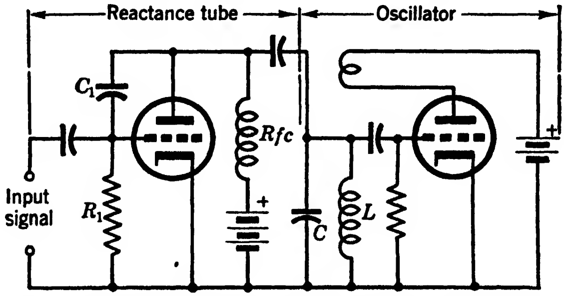


FIG. 12. Circuit of reactance tube and oscillator for frequency modulation.

the resultant reactance of an *LC* tuned circuit is through a reactance tube as illustrated in Fig. 12. The right half of the circuit is the "tickler" coil oscillator described on page 198. The left half of the circuit constitutes a reactance tube component which is in parallel with the *LC* tuned circuit. Values of C1 and R1 are chosen so that the grid voltage across R1 is approximately 90 degrees out of phase with the plate voltage, and the tube therefore acts as a reactive load across the tuned circuit. This reactance is changed at the signal-frequency rate by the input signal at the control grid. In circuits where the frequency drift due to change in components must be kept at the minimum, it is customary to take advantage of push-pull circuits and to exercise particular care in the construction of the component parts. A more elaborate system developed by Armstrong* for frequency modulation uses a crystal oscillator followed by a balanced amplitude modulator, a mixer for obtaining phase modulation, and a series of deviation multiplier stages. The development of special tubes such as the Phasitron and new circuits have increased the application of frequency modulation in radio broadcast service.

* E. H. Armstrong, "A Method of Reducing Disturbances in Radio Signaling by a System of Frequency Modulation," *Proceedings I.R.E.*, Vol. 24, 1936.

Pulse Time Modulation. Pulse time modulation consists essentially in the transmission of signals by pulses of constant amplitude, the instantaneous amplitude of the signal being translated into a variation of time intervals of successive pulses while the rate of this variation corresponds to the instantaneous frequency of the signal. This form of modulation using short pulses of energy may be contrasted with amplitude and frequency modulation where a continuous but varying energy wave is employed.

A concept of pulse time modulation may be obtained by assuming a signal frequency of 10,000 cycles which will divide a second into periods of 100 microseconds. Next, assume that a total period of 4 microseconds is to be allotted for an individual pulse. This individual pulse will be of constant amplitude but may be shifted in position anywhere along this 4-microsecond interval, or the duration of the pulse in this 4-microsecond interval may be varied in accordance with the amplitude of the signal. During the next 96 microseconds no energy is employed for the signal under consideration, but at the end of the period and successive ones a new pulse is produced. The transmitted pulses are picked up by a synchronized receiving device which has a persistence of acceptance (like man's persistence of vision) and does not miss the long off-signal periods.

Advantages of pulse time modulation are twofold. First, since the pulse signals are being transmitted for only a fraction of time, the signal-to-noise ratio is high. Second, for the example assumed above, the 96 microseconds of silent periods may be divided into 24 periods of 4 microseconds, each of which may be utilized for the transmission of another signal or message provided that the transmitting and receiving equipment are properly synchronized. In practice, one of the 25 channels of the preceding pulse time system is used for the synchronizing signal. In passing, it may be noted that the amplitude and the frequency of the signal control the variations of the pulse within its short cycle in a manner somewhat similar to frequency modulation.

Detection Methods and Circuits. Detection or demodulation is the inverse process to modulation. Thus detection involves the extraction of the modulation signal from the modulated carrier. Most detection circuits utilize rectification and the unilateral characteristics of vacuum tubes and other nonlinear devices as the basis of operation. Detection methods for amplitude-modulated carriers are divided into two classes, *linear* and *square law.* In linear detection the current flows in one direction, falls to zero, and exists in pulses. In square law detection the current in the detecting element is unidirectional,

pulsating, and does not fall to zero. The simplest form of linear detection is illustrated in the diode detector circuit of Fig. 13. Here a filter * consisting of a capacitor and resistor in parallel serves to detect the signal contained in a modulated carrier. The action of the RC

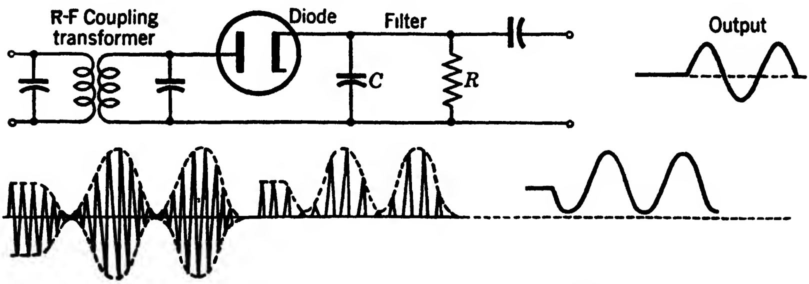


Fig. 13. Typical circuit and components in diode detection.

combination can be followed easily by a study of Fig. 14. First the unilateral characteristic of the diode will serve to cut off on all negative voltage pulses (shown dotted in figure). Beginning with a low-voltage positive pulse on the left, the tube will conduct when the plate is at a higher potential than the cathode. The electron flow (up)

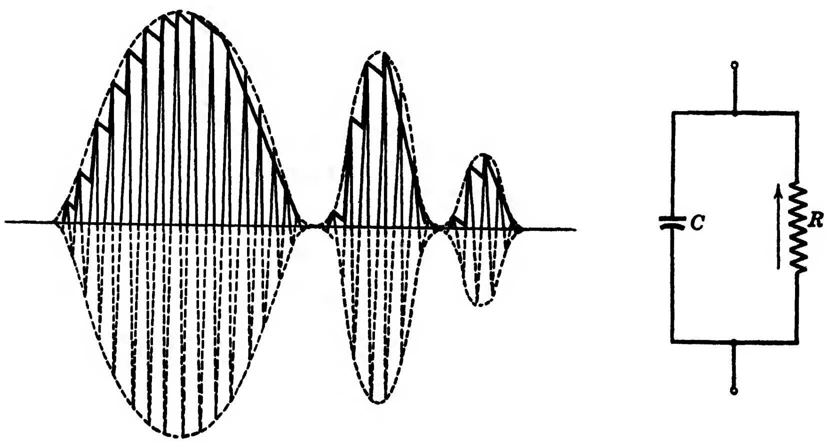


Fig. 14. Rectified voltage across RC in diode detection.

through R will cause the top terminal of the resistor to be at a higher (positive) potential, and the condenser C charges to this potential. When the impressed voltage reaches its first peak and starts to fall, the tube will cut off because the capacitor will hold its positive poten-

* Filters are treated on pp. 347–353.

tial for a time and cause the potential on the plate to be lower than on the cathode. While the impressed wave of voltage passes through the major part of the succeeding cycle (down to zero, negative maximum, up to zero, etc.), the capacitor C discharges slowly through resistor R in accord with its time constant and the voltage across RC falls producing the first tooth in the saw-tooth voltage wave. When the next positive wave of impressed voltage causes the plate to have a higher voltage than the cathode, rectification takes place and the condenser is charged again to a higher potential. As the process is followed it is evident that the potential across the RC detector combination follows the jagged form shown in Fig. 14. This form is the envelope of the modulated carrier which under amplitude modulation represents the original modulating signal. The changing form of the waves in the different parts of the detecting circuit is shown in Fig. 13.

The detection efficiency of the diode detector is defined as the ratio of the voltage developed across the resistor to the voltage of the impressed modulated wave. High detection efficiency may be attained by keeping the value of R high with respect to internal resistance of the diode r_p. If R/r_p is of the order of 20 to 100, detection efficiency will lie in the range of 70 to 90 per cent.* Also, when the ratio of the a-c load impedance to d-c load resistance is equal to or greater than the modulation factor, distortion will be low. Since all diodes have some curvature near the beginning of their i_b-e_b characteristics, it is preferable to work with rather large signal voltages. The time constant of RC is very important. This constant should be large compared to the time of one cycle of the rf carrier in order to keep the steps in the output wave small. Yet the time constant must be small compared to the period of the highest modulation frequency so that the charge on C can decay rapidly enough to follow the downward swings in modulation. Typical values are 250,000 ohms for R and 100 to 250 $\mu\mu f$ for C.

In the simple RC diode detector circuit some unmodulated carrier may be present across R and the detected potential is a variable d-c voltage. It may be desired to amplify the detected signal, or use it in a manner where the carrier and d-c components are objectionable. To remove these components from the detected signal and to improve the characteristics of the detector, RC filters may be added to the simple circuit as shown in Fig. 15. To eliminate the carrier the R_2C_2 combination may be added, and to remove the d-c component the C_3R_3 component is desirable.

* Terman, *Radio Engineers Handbook,* McGraw-Hill, 1943, p. 554.

The simple diode detector circuit can be modified to give automatic volume control as shown in Fig. 16. The magnitude of the rectified current will be proportional to the volume of the incoming signal and the amplitude of the modulated voltage impressed. The electron current through R causes point X to be negative with respect to ground and the magnitude of this negative potential varies with the volume. This negative potential at X is used as the grid bias for an earlier amplifier stage after passing through the R_cC_c filter. The time constant of the filter R_cC_c is very long compared to the frequency of the audio signal and long compared to the normal changes of inflection of the human voice, but short with respect to the period of signal fading resulting from changes in the Heaviside layer. Thus the proper adjustment of an automatic volume control will vary the gain of a preceding stage in an amplifier to compensate for changes in incoming signal strength.

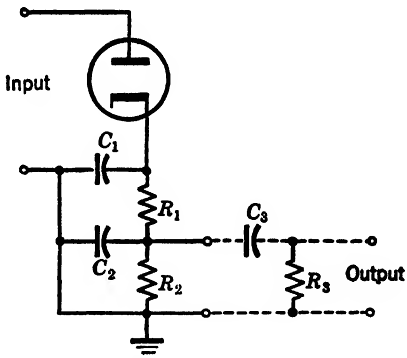


FIG. 15. Improved RC detector circuit to remove carrier and d-c component.

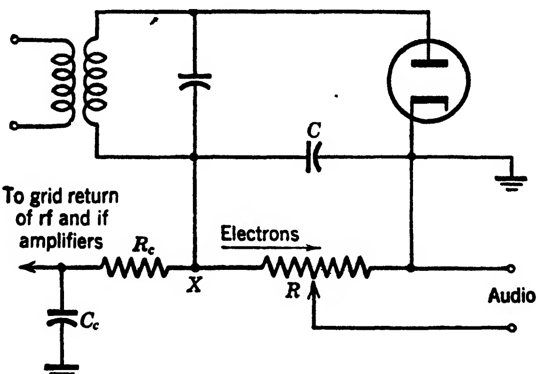


FIG. 16. Automatic volume control circuit.

Diode detectors are widely used because (1) the circuit is simple, (2) distortion is low, (3) large signals can be handled without overload, and (4) the d-c component of the rectified signal can be used to provide automatic volume control.

The new silicon and germanium crystals described in Chapter XIII have rectifying properties similar to the diode, and they are now being employed rather generally as detectors in high- and ultra-high-frequency circuits.

A grid-leak detector is approximately equivalent to a diode detector plus one stage of amplification. Its typical circuit is given in Fig. 17 where cathode-grid rectification takes the place of cathode-plate rectification in the diode circuit. The R_gC_g component in the grid circuit has a rectified d-c component of the form of the modulating signal.

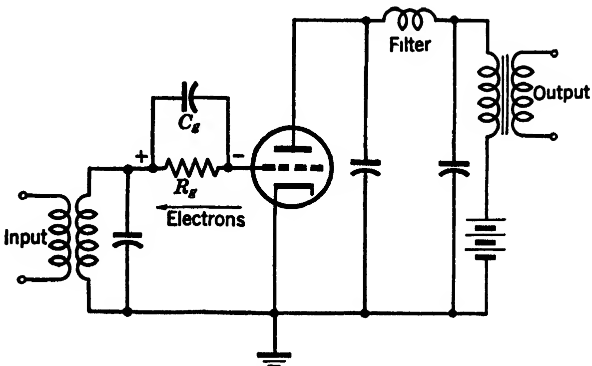


FIG. 17. Grid-leak detector circuit.

This voltage constitutes the grid signal voltage which controls the amplified output in the plate circuit. The LC π type filter in the plate circuit by-passes the carrier frequencies and forces the low signal frequency into the transformer output circuit.

The grid-leak detector has the advantage of high sensitivity, but the amount of distortion generated at high modulation levels is objectionable. Large rf signals cause excessive distortion because the grid is driven too far negative for proper amplifier operation.

In order to obtain even greater sensitivity and improved selectivity, the grid-leak detector may be made regenerative. The circuit looks very much like a feedback oscillator with an R_gC_g in the grid path. However, the grid leak R_g is much higher and the coupling between L_1 and L_2 is adjusted just below the point of oscillation. Plate circuit energy is fed back to cancel the losses of the grid circuit, the Q of the resonant tank being thereby greatly increased and the selectivity improved. Regenerative detectors may produce a large amount of distortion and are difficult to adjust since the amount of feedback coupling required varies with the tuning of the grid circuit. Greatest sen-

sitivity is obtained when the regenerative detector is adjusted so that it is just below the point of oscillation. This point is difficult to maintain, and so occasionally a scheme known as super-regeneration is employed. Here an alternating voltage is impressed on the plate or grid circuit in such a manner as to block or "quench" any oscillation in the detector. The quenching frequency must be high compared to

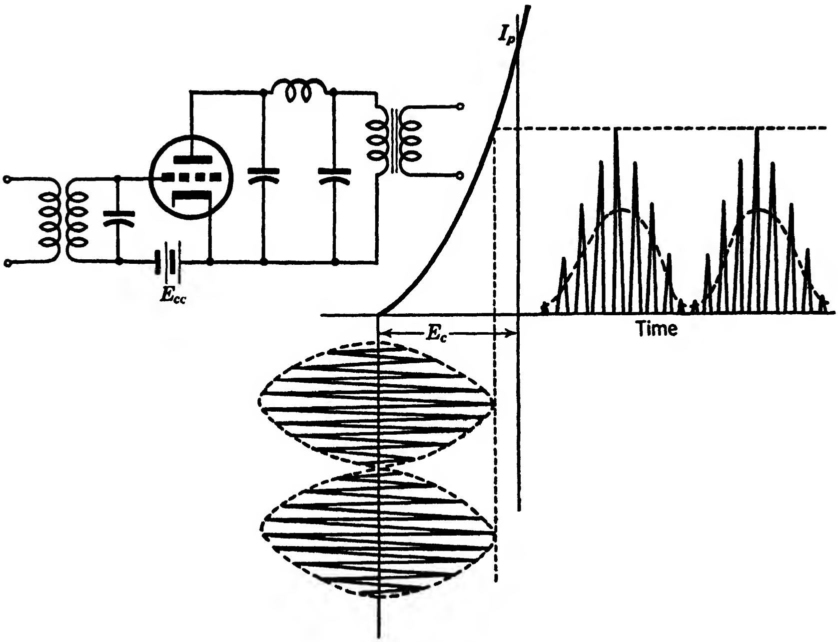


Fig. 18. Plate detection and circuit.

the audio signal, yet low when compared to the carrier frequency. Feedback can then be increased so that the detector continually breaks into oscillation. This gives much greater sensitivity and can be obtained by any other type of detector although the adjustments are rather critical and the selectivity is poorer than the straight regenerative circuit. Super-regenerative detectors are useful at high radio frequencies where rf amplifiers are ineffective.

The substitution of a fixed grid-bias supply E_{co} for the voltage drop across RC in Fig. 17 gives a form of circuit which may be used for plate detection. In this circuit, shown in Fig. 18, the grid is biased to the cutoff (Class B) so that only the positive loops of plate current appear. Since operation is over a curved portion of the dynamic characteristic, square law detection takes place. The average value

of the plate current represents the modulating signal which may be removed by a suitable plate load circuit. This detector will deliver a large output voltage with moderate sensitivity and distortion. It has the advantage of presenting a high impedance to the tuned circuit since the grid is always negative. No rectified voltage is available for use in automatic volume control.

The principle of detection may be employed for the construction of a *vacuum-tube voltmeter.* One simple circuit for a vacuum-tube voltmeter is suggested in Fig. 19. The tube is biased to cutoff by a fixed bias or other suitable circuit for producing such bias. A milliammeter is connected at A to indicate the changes in the grid supply voltage. Reference to Fig. 14, Chapter V, shows how a varying signal on the grid causes a rise in the average value of the plate current due to the nonlinear dynamic characteristic. With the grid biased to zero the plate current will be zero for zero signal voltage, but the average value of plate current will rise with the amplitude of any applied a-c signal. The milliammeter A must be calibrated by using known values of applied a-c voltages. Most vacuum-tube voltmeters present a very high input impedance and require an infinitesimal amount of current for operation. These properties make them very useful for measurements in circuits where the impedance relations should not be disturbed and where the power drain should be infinitesimal.

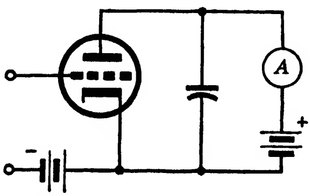


FIG. 19. Simple vacuum-tube voltmeter circuit.

In order to recover the audio signal from the FM carrier, it is necessary to convert the frequency deviation back to amplitude variation. This can be done very simply by adjusting a tuned radio-frequency amplifier so that the mean carrier frequency falls on one side of the selectivity curve (see Fig. 5, Chapter V). As the frequency varies, the output from the amplifier will vary in amplitude. Such a detector is likely to produce large amounts of distortion except for very small deviations because the skirts of the selectivity curve are not linear over a very wide frequency range. Furthermore, such a detector will respond to any amplitude modulation present on the carrier and produce a component known as noise. A better detector uses a balanced diode circuit called a discriminator. The explanation of this device is beyond the scope of this book. In order to realize the full benefits which FM offers in regard to elimination of interference and noise, it is necessary to use one or more limiter stages preceding the

FM detector. Limiters are simply intermediate-frequency amplifiers (explained in next article) operated at their overload point so that any change in input amplitude will result in very little change in output. Properly designed limiters will give constant output signals with a variation of several hundred to one of input.

As an aid to obtaining better signal-to-noise ratio in the transmission of radio programs, "pre-emphasis" has been widely applied, particularly in connection with FM broadcasting. Pre-emphasis is based upon the fact that the amplitude of the various frequency components in an average radio program decreases as the frequency increases. Therefore, a network is inserted in the audio circuits of the transmitter which increases the relative amplitude of the higher frequency so that the 15,000-cycle components are 20 db greater in amplitude than the 1000-cycle components. This has the effect of increasing the per cent of modulation for the higher frequencies and results in an improved signal-to-noise ratio. In the receiver it is necessary, of course, to have a complementary network which will reduce the relative amplitude of the high frequencies and restore them to their natural proportions. The benefits of pre-emphasis can be realized as well with AM as with FM transmission. However, so many AM receivers have been built without the required correcting networks that it is impracticable to use pre-emphasis in the present AM broadcasting service.

Heterodyne Systems. The heterodyne or "beat-frequency" system is a useful method of creating a desired frequency by employing the process of detection. When two alternating voltages of different frequency are mixed or superimposed, the amplitude of the resulting wave varies at a rate corresponding to the difference in frequencies. If this resulting wave is impressed across a suitable detector the output will be the difference or beat frequency of the original waves. As an example, an rf wave with a frequency of 5000 kilocycles may be changed to 500-kilocycles by combining it with a wave having a frequency of either 4500 or 5500 kilocycles. The combined signal is impressed on a heterodyne detector with a load output tuned to 500 kilocycles.

The heterodyne principle has many applications. For laboratory purposes a signal generator may be constructed by coupling two rf oscillators to a detector. If one oscillator has a fixed frequency f and the other is variable with a frequency $(f + \Delta f)$, a wide range of beat or difference frequency Δf will be available in the output. Since a high-frequency oscillator is easier to construct and control, this application is sometimes preferable to a simple oscillator capable of producing the

desired frequency range. When two frequencies are mixed in the heterodyne system, if one signal has a varying frequency, the beat frequency, though of lower frequency value, will contain the variations. This makes it possible in the superheterodyne radio receiver to mix a signal voltage from a local oscillator with the received modulated signal and convert to a lower or intermediate frequency band. The lower or intermediate frequency can be amplified by a high-gain "intermediate amplifier" tuned for a fixed frequency band. This intermediate stage in the receiver makes it possible to amplify and detect signals from all broadcast bands with the maximum of efficiency and the minimum of distortion compared to an arrangement whereby the amplifier and detection stages must handle a wide range of frequencies.

If an intermediate amplifier stage is overloaded by the applied a-c grid signal, the output waves in the plate circuit will have a constant amplitude. When an amplifier is operated in this manner it is termed a *limiter*.

Copper Oxide Varistors as Modulators and Demodulators. Nonlinear devices other than vacuum tubes may be used in the process of

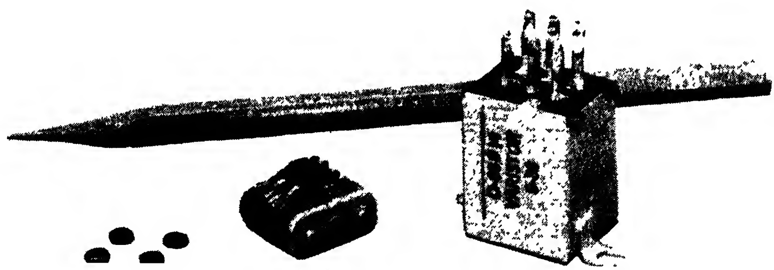


FIG. 20. Varistor and parts used in modulators and demodulators. (Courtesy Bell Telephone Laboratories.)

modulation. One such device is the copper oxide rectifier described on pages 318–320. Four tiny copper oxide units possessing unilateral conductivity are connected in a bridge circuit and mounted in a small container as illustrated in Fig. 20. These assemblies, known as Varistors, are widely used as modulators and demodulators on carrier telephone systems. One basic circuit for this application is shown in Fig. 21 where the voice signal is applied across one pair of opposite

points on the bridge and the carrier signal to the other pair of terminals. The arrows for the CuO unit indicates the conventional direc-

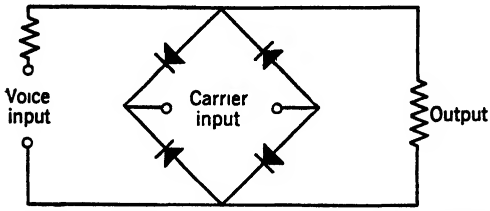


FIG. 21. Bridge modulator circuit using Varistors.

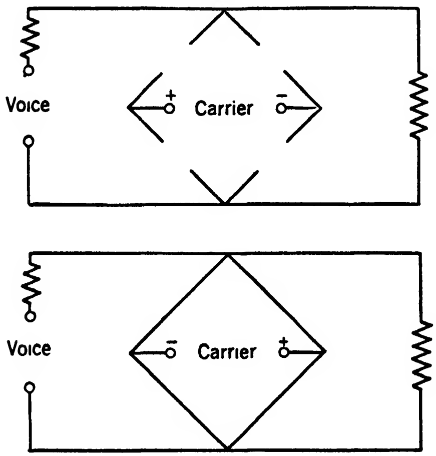


FIG. 22. Operation of Varistors in modulator circuit.

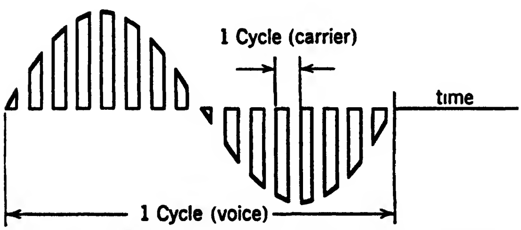


FIG. 23. Modulated wave produced by Varistors.

tion of current for each rectifier unit (resistance is low in this direction and very high in the opposite direction). The applied carrier voltage is high and causes the bridge to act as an open circuit and then as a closed circuit for the voice signal, as illustrated in Fig. 22.

The action is such that the carrier voltage causes the bridge to act as a "chopper" for the applied signal, resulting in a transmitted wave as shown in Fig. 23. An analysis of this current wave will show that its principal components are the signal frequency and the upper and lower side bands of the carrier frequency. Other components present are of low magnitude. In carrier telephone systems one side band is transmitted and all others eliminated by a filter. At the receiving end demodulation is accomplished by applying the transmitted side band as the voice signal in the circuit of Fig. 21. Here the output of the modulator contains the original voice signal as one of its components. This component is selected by a filter and amplified for transmission to the local subscriber's station. Other circuit arrangements of Varistors are employed for carrier service though the underlying principle is similar in all of them.

Chapter XI

ELECTRICAL CONDUCTION IN GASES

For a long time gases were supposed to be perfect insulators. Dry air and other gases seemed to offer a high opposition to the flow of electric current. About 1900 J. J. Thomson performed experiments which showed that gaseous ions serve as carriers for the electric current.

The conduction of electric current in gases is not easily predictable since it depends on many variables. The resulting conduction may vary with the gas employed, the gas pressure, the potential between electrodes, the electrode material, the shape of electrodes, the distance between electrodes, the shape of the enclosing medium, and other factors. Conduction in gases may be attained with either hot or cold cathodes, but the action will be different in each case. All this is in contrast with the more simple theory of conduction of electricity in metals and in vacuums which has been considered in the preceding parts of this text. Some aspects of the phenomenon of gaseous conduction will be covered in this chapter to aid in the understanding of the action of gaseous electron tubes which are widely employed in the field of electronics.

Kinetic Theory of Gases. The molecules of a gas are in a constant state of motion similar to that of molecules in liquids and solids. Those in gases enjoy a greater freedom of movement so that what is termed gas pressure is really the result of multiple impacts of gas molecules upon the walls of the restraining enclosure. The simple kinetic theory of gases assumes the molecules to be small spheres which collide with each other in the course of their constant motion. The distance a molecule moves before it collides with another is called its free path. Obviously, the length of paths vary greatly, some being relatively short and others long. A study of the distribution of these paths will give a *mean length of free path* or average path which is of importance in the theory of electrical conduction in gases. The mean length of free path depends on the gas pressure and rises in magnitude as the pressure

falls and the molecules are farther apart. The mean length of free path will also depend upon the size of the molecules. Mean length of free path of gaseous molecules is of the order of 0.02 to 0.2 millimeter for a pressure of 1 millimeter of mercury.

An electron projected in a gas will likewise collide with the molecules present. After its first collision it will bounce off in a new direction until it experiences a second collision, and then in a new direction for a third collision, and so on. In this manner it will travel in a zigzag path through the gas. The distance between collisions is the length of free path for the electron, and the average length of these paths is the *mean length of free path* for the electron. Since the electron has a smaller mass and a smaller size than the molecules of gas, it will experience fewer collisions and will have a longer length of free path (approximately six times that of the molecule).

The molecules of the noble gases and mercury vapor consist of one atom. These atoms possess kinetic energy because of their thermal agitation, as suggested. In addition to this kinetic energy, these atoms may be given potential energy by displacements of electrons in their atomic structure. In the normal state the electrons in the atom exist in certain orbits, rings, or energy levels. If one or more electrons in an outer orbit are disturbed by the addition of energy, they may be moved out of their normal orbit or energy level into a higher level or they may even be removed from the atom. In this process the atom acquires new energy. This new energy may be released or radiated very quickly with a return to a normal state, or it may be retained for a short time before release. Thus a gas atom is capable of receiving, transporting, and releasing energy through rapid changes in its atomic structure. This property combined with its kinetic energy plays a very important part in electrical conduction in gases.

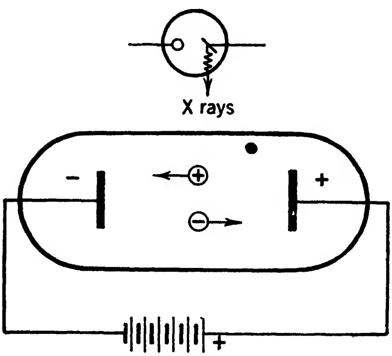


FIG. 1. Simple gaseous conduction.

Gaseous Conduction. The term gaseous conduction implies the conduction of current by a gas itself and within itself. Such conduction is simple to understand. The gas-filled tube of Fig. 1 contains two cold electrodes held at a difference of potential by a battery. Assume that one gas atom has been ionized to produce one positive

ion and one electron through the action of X rays. The electron will be attracted to the positive electrode and will enter the circuit leading to the battery. The positive ion will be attracted to the negative electrode and upon arrival will seize an electron from the plate and unite with it to form a neutral atom of gas. This simple dual action has removed one electron from the negative electrode and supplied one electron to the positive electrode. This transfer constitutes an electric current. Multiply this single transfer by millions and simple gaseous conduction results.

The ions taking part in gaseous conduction must be produced by some secondary action. Often such secondary action involves one or more forms of electron emission. When this happens the emitted electrons join the ion movement, and the total current through the gas consists of the pure gaseous conduction plus an electron current. Thus the electric conduction in gases usually involves two or more processes of electron transfer.

Methods of Producing Ions. Gases may be ionized (1) by thermal action, (2) by electromagnetic radiation, and (3) by collision with particles. Gases in a flame become ionized by thermal action. X rays, gamma rays, cosmic rays, and other forms of electromagnetic radiation such as ultraviolet light, have the power of ionizing a gas. Alpha particles and beta particles released from radioactive materials collide with gas particles and leave an ionized path. Much of our early and later scientific study has utilized the ionizing property of flying particles and the energy of radiation. In gaseous electronic devices, ionization by collision with electrons plays a very vital role.

Ionization by collision with electrons is produced in gases by the accelerating force of an electric field. The primary electrons for the ionizing process are generally released by thermionic emission or photoemission. To get a concept of ionization by collision, assume for the gaseous tube of Fig. 2 that the gas pressure and the electric field are of suitable value for ionization to occur. Electrons emitted from the negative hot cathode will be accelerated by the electric field toward the positive anode, resulting in collisions with atoms of gas while in flight. A certain collision with a Bohr atom shown in the figure results in ionization of that atom. Here the approaching primary electron is aimed at an electron in the outer orbit of the atom. This orbital electron is moving to the right, and the collision with the primary electron also moving to the right will increase its speed so that it will break away from the atom. The orbital electron carries a negative charge

$(-e)$ and the atom becomes a positive ion with a charge equal to $(+e)$ in magnitude. The primary impinging electron glances off at some angle as shown and continues its transit until it reaches the positive anode. During the ionizing collision the primary electron gives up a definite amount of energy to the atom. Hence the resulting positive ion acquires and carries a unit of energy which will be released again whenever the ion reverts to a normal atom.

Each electron passing between the electrodes of the tube of Fig. 2 under the influence of the electric field suffers many collisions with the atoms in the intervening space. Only a few (less than one per cent)

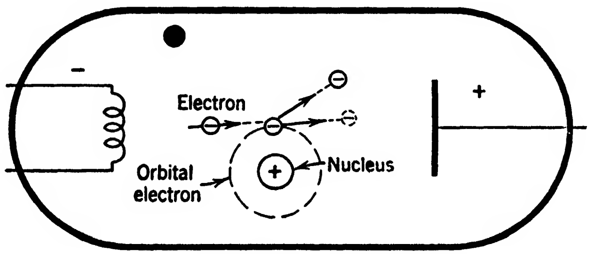


FIG. 2. Ionization by collision.

of these collisions result in ionization. This follows because the electron may not possess sufficient energy to produce ionization or because the nature of the collision is not favorable for ionization.

The preceding discussion might lead one to assume that the collisions of molecules and the impacts between flying electrons and atoms were actual physical contacts like those of a baseball bat hitting a ball. Such is probably not true. According to the concept of atomic structure, collisions and impacts may become the interaction of the charges on electrons, atomic nuclei, and molecules. As one molecule of gas approaches another, the electrons in their outer orbits exert a repulsion for each other, and when they approach close enough this force of repulsion may become sufficient to cause the molecules to stop and then bound away from each other. In like manner, when a high-speed electron approaches an atom, if the electron in the outer orbit happens to be in the direct path of the projected electron, the force of repulsion between the two may carry the one free from the atom and produce ionization. If one recalls that the ratio of the charge on the electron to its mass for moderate velocities is 1.76×10^{11} (mks system), it is obvious that the force exerted by the charges vastly outweighs the importance of the insignificant mass. It should be noted that, since atoms

are nebulous, electrons and even positive ions may pass through atoms provided they do not move on a line through the electrons or the nucleus.

The intensity of ionization in a gas depends on several factors such as applied potential, size of plates, distance between plates, and gas pressure. If two electrodes are placed in a gas at atmospheric pressure, a low difference of potential would sweep any ions present out of the gas. However, it would require a very high voltage to produce any ionization. This follows because the gas molecules are packed closely and any electron moving under the existing potential gradient could not be accelerated sufficiently before a collision to produce ionization. With a lowering of the gas pressure, the mean free path of the molecules increases, the electron moves farther, gaining more speed before collision, and ionization occurs more readily. The potential necessary to start ionization under any condition of gas pressure, etc., is often called the ionizing potential. As the pressure of the gas is reduced the ionizing potential will continue to decrease. Assume a moderate but fixed potential between the plates and then decrease the gas pressure by gradual steps. At a high gas pressure few if any electrons will ever attain sufficient speed to produce ionization. As the pressure is reduced the ionizing potential for that condition is reached and a current flows. A further reduction in pressure allows electrons to move farther before collision, and hence ionization becomes more intense. This increase of the intensity of ionization will continue with the decrease of pressure until the maximum number of the gas atoms is taking part in the ionization. Then any further decrease in gas pressure will reduce the number of atoms present so that the actual current conducted by gas ions will decrease progressively with pressure, ultimately approaching zero in a very high vacuum.

The term ionizing potential is frequently used rather loosely. As commonly applied, it refers to that potential which must be applied between two electrodes to produce ionization for a given gas, pressure, temperature, and electrode spacing. In a specific sense, one may think of the ionizing potential as the voltage through which an electron must fall to produce ionization. Following this concept, each gaseous element requires a definite energy in electron-volts to produce excitation and ionization. For example, helium has a minimum excitation potential of approximately 20 volts and an ionizing potential of approximately 25 volts. For mercury vapor the approximate excitation value is 4.7 volts with an ionizing value of 10.4 volts. The distance through which an electron can fall before collision depends upon the

gas pressure and temperature. Obviously, these latter factors govern the potential gradient and the difference of potential between electrodes for producing ionization.

The collision of the electrons with gas atoms may be divided into four groups. First, many collisions result in a mere "bouncing off" or change of direction of the electron where little or no energy is imparted to the atom. These are known as *elastic collisions* and constitute the majority of all impacts. Second, some collisions occur where the energy of the electron is insufficient to produce ionization but is great enough to move an outer electron out of its orbit or energy level. Here the atom has been given some energy and is said to be *excited.* This excited state of the atom is a very transient one and the atom releases the acquired energy almost immediately. The energy is released in the form of electromagnetic waves. This wave energy radiation may be *visible light* and as such is very important in some types of electric-lighting sources. In a third group of collisions, the electron is able to impart energy to the atom so that it is retained for a short period of time. While this energy is retained, the atom is said to be in a metastable state. A metastable atom may receive additional energy from a second collision which will be enough to produce ionization. It should be noted that the excited atom and the metastable atom retain all electrons and hence do not carry a charge and are not affected by an electric field. The fourth group of collisions is the *ionizing collision.* If the conditions are favorable for ionization to occur, all four groups of collision will be taking place.

Gases may be ionized by the collision of positive ions with atoms. However, the positive ions present in gaseous tubes are not very effective ionizing agents. Their large mass relative to the electron prevents rapid acceleration in an electric field and their larger size results in more collisions with molecules, so they have little opportunity to acquire velocities sufficient for ionizing collisions. It is well to remember that in the simplest gas, hydrogen, tne positive ion (a proton) is 1840 times as heavy as the electron. In the more complex gaseous elements used in commercial tubes, the positive ion has a mass several thousand times as large as the electron (see Table 1, page 7).

Deionization. Deionization is the reverse process of ionization. It is effected by a recombination of positive ions and electrons to form neutral normal atoms. Such recombinations are accompanied by a release of the energy required for ionization and imprisoned in the positive ion during its existence. Recombination may take place (1) inside the volume of gas, (2) along the walls of the gas-enclosing

chamber, or (3) at an electrode under the attraction of an electric field. Recombinations do not take place readily inside the gas volume if an electric field is present because positive ions and electrons will have high relative velocities in opposite directions. If the conditions are favorable for the formation of negative ions (atom plus an electron), recombination may take place readily since the negative ion with larger mass moves more slowly than an electron, even in the presence of an electric field. Also the negative ion gives up its extra electron readily to a positive ion. Positive ions which diffuse to surface walls obtain electrons readily for deionization. Electrodes having negative potentials attract positive ions and supply electrons for recombinations. The energy released during deionization may be in the form of heat and may aid electron emission at the cathode. The preceding statements on recombinations apply to low-pressure discharge conditions but not to high-pressure discharges where mean free paths are negligible compared to electrode spacing.

Stages of Electrical Conduction in Gases. Electrical conduction in a gas between cold electrodes may pass through and exist in three successive stages. These stages are called the Townsend discharge, the glow discharge, and the arc. The term discharge arose from early experiments wherein gaseous conduction served to discharge electricity stored in condensers. The Townsend discharge can be explained by the circuit and curve of Fig. 3. Two cold plates are placed in the chamber C containing gas under reduced pressure. A variable voltage from source E_x is placed across the plates. Beginning with X at zero or a, the voltage is raised as X moves to position b with a current flow in the circuit as shown by the curve $OPQR$. While no ionizing agent is provided within the chamber, a small number of ions will be present because of cosmic rays and other electromagnetic radiations which exist outside of the chamber or because of radioactive substances which may be present on the inside or outside. The small number of ions present and being formed constantly will be attracted to the electrodes and will constitute a *very minute current* which will rise along the line OP under the influence of a weak field. The voltage at P is sufficient to sweep all these ions out of space as rapidly as they are formed, and from point P to Q there will be no change of current. At point Q the ionizing potential for the pressure and gas used is reached. Hence a few ions will be produced and their transfer will add to the very minute current present. Now, as the voltage is raised above the ionizing potential, more primary electrons will produce ionization.

Also, as the potential rises, two things may happen. First, an occasional primary electron may have more than one ionizing collision while in transit, and, second, a few electrons formed by collisions may in turn effect an ionizing collision. In this manner, the process of ionization becomes accumulative and the gaseous conduction current rises rapidly following an exponential type of curve. Ultimately, as the point X is moved to the right, the current tends to rise without

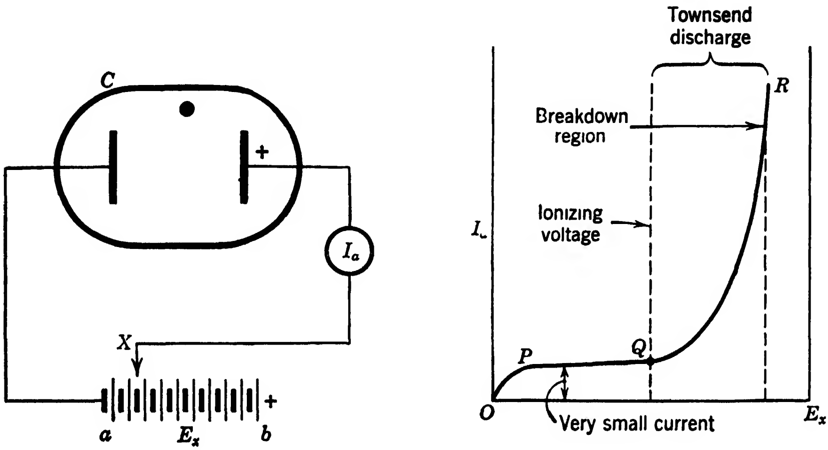


FIG. 3. Circuit and phenomena for Townsend discharge.

apparent limit until a breakdown region is reached. The breakdown region is marked by two occurrences. First, the operation of the tube becomes self-sustaining, and, second, the discharge passes into a second stage or form of discharge. Up to this point the discharge has depended upon the production of some ions by an external or separate agent, such as electromagnetic radiations. If this external action had been stopped, the discharge current would have stopped likewise. At the breakdown point a stage has been reached where the products of ionization (electrons and positive ions) will reproduce sufficient electrons and ions to keep the discharge continuous. This action is accomplished by electron emission at the cathode resulting from bombardment by the positive ions attracted to it plus the energy released during recombination. One factor in the breakdown of the Townsend discharge is the change in the potential distribution between the electrodes resulting from the rising positive space charge which increases the potential gradient near the cathode. After reaching the breakdown point, the Townsend discharge may pass into other types of discharge

such as the corona, the glow, and the arc. The term breakdown potential is known also as sparking potential and ignition potential.

The breakdown voltage E_b in the Townsend discharge depends upon the geometry of the electrodes, their work function, the gas employed, the gas pressure, and the distance between the electrodes. In order to generalize on the phenomenon, a particular gas and electrode material may be chosen and the cold electrodes may be assumed to be parallel plates, thus giving a uniform electric field if space-charge effects are ignored. If the electrodes are held at a fixed spacing and the gas pressure is increased from zero to X, the trend of the breakdown voltage E_b varies as shown in Fig. 4. Here the mean free path of the electron will control the intensity of ionization and, in turn, the breakdown point, as previously discussed on page 249. Next, assume the gas pressure is held constant and vary the spacing between the electrodes from zero to some considerable distance X. The measured breakdown voltage will again follow the same trend, as shown in Fig. 4. Here also, the length of free path of the electron determines the characteristic curve. At short spacings the electrons experience few collisions and the necessary breakdown voltage must be very high. For large spacing the distance may equal several lengths of free path, and again E_b must be high. The optimum condition for breakdown is likely to occur at a spacing approximating the mean free length of path for the conditions assumed. This characteristic variation of breakdown voltage with electrode spacing has been utilized in the design of cold-cathode rectifiers and grid-glow tubes to be covered in the following chapter. Thus a spacing corresponding to point b on Fig. 4 requires a relatively low breakdown voltage, whereas a short spacing such as point a will permit three times such potential before conduction occurs.

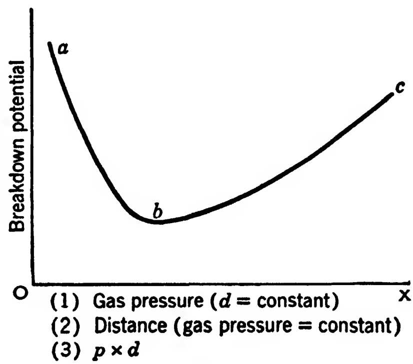


FIG. 4

The preceding discussion discloses that the trend of change of breakdown voltage with variation of gas pressure for a constant electrode spacing is the same as for the inverse condition, namely, a variation of electrode spacing for a constant gas pressure. From this disclosure it is obvious that a similar trend will follow for a variation in the

product of gas pressure times distance, or $p \times d$. This relationship is known as *Paschen's law.* Paschen's law states that, for parallel plane electrodes in a given gas at a given temperature, the breakdown voltage is a function of the product of the pressure and the electrode separation. To understand this law one should remember that under the conditions assumed the mean free path is inversely proportional to the pressure, so that the ratio between the spacing and the mean free path remains

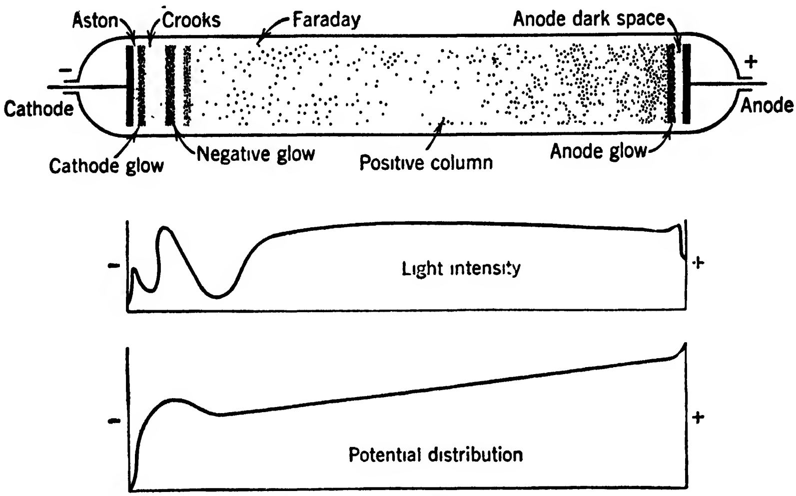


FIG. 5. Distribution of quantities in a long gaseous column with a cold cathode. (Reprinted with permission from L. B. Loeb, *Fundamental Processes of Electrical Discharge in Gases,* John Wiley & Sons, 1939, p. 566, Fig. 269.)

the same. Thus the number of free paths between the electrodes remains the same for a particular product of $p \times d$.

The *glow discharge* is a self-maintaining discharge which is marked by a luminous column, a low current density, and a high voltage drop between electrodes. This discharge gets its name from its soft luminous effect exemplified in the familiar neon sign. Some of the properties of the glow discharge for a long tubular chamber under an applied d-c potential are indicated in Fig. 5. Most of the light comes from the long positive column with supplementary bright areas at the cathode glow, the anode glow, and the negative glow points. The potential distribution shows a rapid rise of voltage close to cathode, then a gradual rise throughout the long positive column, with a final spurt or rise in potential near the anode. This form of potential distribution may be explained by the slow mobility of the positive ions in the long positive

column. The ions form a positive space charge which moves down to the cathode and gives a form of virtual anode close to the cathode. The rapid rise in voltage beginning at the cathode is equal to the breakdown voltage for the particular pressure and gas used. This voltage is several times the ionizing potential and it produces primary electrons and ions in sufficient numbers to make the discharge self-maintaining. This rise in voltage, often called the cathode fall of potential, depends upon the material used in the electrodes, upon the kind of gas, and upon the gas pressure in the tube. This voltage lies within the range of 50 to 300 volts. The positive ions are repelled by the positive potential at the anode so that only electrons exist in a narrow region called the anode dark space. In this region a more rapid rise of potential (stronger electric field) is needed to withdraw the electrons from the positive column.

When the current in the glow discharge is small the cathode glow covers a small area of the cathode, and the area increases in size with the magnitude of the current. Since the tube voltage is nearly constant with current variation, it is necessary to limit or control the current by some external means. In neon signs this control is built into the transformer that supplies the voltage. The constant voltage characteristic of the glow discharge is utilized in voltage-regulator tubes to be described later.

If the current density in the glow discharge is permitted to pass a certain maximum value for a particular gas, the character of the discharge changes rapidly. The increase of current will be accompanied by a rise of voltage drop between the cathode and anode and the glow will rise in intensity. This new state is called the abnormal glow. A further increase of current causes an acceleration in the brightness of the glow and a rise in voltage until the discharge suddenly changes to a new stage called the *arc*. The important characteristic of the arc is shown in Fig. 6. Here the voltage across the arc varies inversely with the current through the arc. This is a negative resistance characteristic which tends to damage any equipment associated with the arc. Hence the arc calls for a stabilizing element in the external circuit which will limit the current.

The phenomenon of the electric arc discharge is not well understood. The arc itself is made up of positive ions, electrons, excited atoms, and metastable atoms. The high current density in the arc results in very high temperatures. This high temperature may influence the arc phenomenon in several ways. First, the high temperature may increase electron emission from the cathode and thus increase the current.

Second, the temperature may be sufficient to produce ionization by thermal agitation. Third, the high temperature reduces the ionizing energy necessary to produce ionization, thus permitting an increase of ionization by collision. All these processes may act accumulatively to produce intense electrical conduction even with exceedingly short electrode spacing. If the arc takes place in the presence of oxygen the electrodes will be burned.

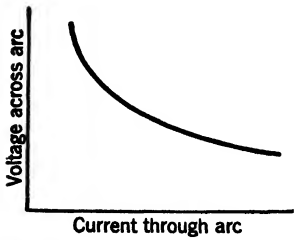


FIG. 6. Characteristic of the electric arc.

The three stages of gaseous discharge with cold cathodes are briefly summarized in Fig. 7. In the Townsend discharge initial conduction is produced by some ionizing agent and a minute current exists under a rising potential until the ionizing potential is reached. After reaching this point the conduction current rises rapidly with voltage, and the potential distribution shifts to bring about a new stage wherein the discharge is self-sustaining. In the new stage (the glow discharge) the potential drop across the tube is constant while the current rises until the cathode glow covers the cathode surface. A further rise in current flow results in a rise in cathode-anode drop, a rise in luminous brightness, and a change to the arc discharge. The arc discharge is marked by an intense luminosity and a fall of potential with a rise in

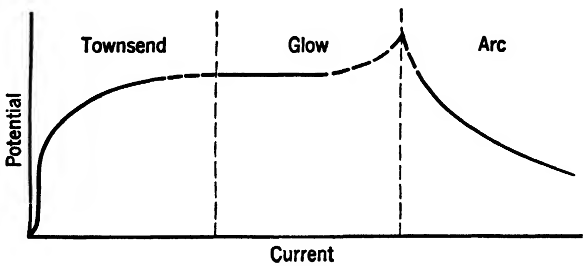


FIG. 7. Summary of the stages of discharge between cold electrodes in a low-pressure gas.

current. Thus the transitions in the three stages are marked by (1) a rise of current and a rise in potential, (2) a rise of current and a constant potential, and (3) a rise of current and a fall of potential.

In all forms of electrical conduction in gases the positive ions play the vital part. The large mass and slow mobility of these ions cause them to form a dense concentration in the space where conduction is

taking place. Langmuir has termed such concentrations or positive space-charge regions the *plasma*. The plasma has boundaries on all sides. For example, in Fig. 5, the end boundaries occur at the negative glow and the anode glow. The other boundary is the outside of the column. This outer boundary may be the wall of the tube but in general the outside boundary is a layer or sheath. This *sheath* is a region consisting of ions, metastable atoms, and normal atoms. The sheath differs from the plasma in that the ions are in a general state of diffusion and do not have a coordinated drift as a part of the current. The sheath can be likened to a thin layer of water on the inside of a pipe having a rough inner surface. This outer skin of water may be nearly stationary and yet serve as a boundary for the water flowing through the pipes. In gaseous conduction tubes, sheaths may represent the boundaries for electrodes as well as enclosing surfaces.

The preceding discussion of electric conduction in gases has assumed the use of a cold cathode, whereas most of the gaseous tubes in commercial use employ hot cathodes. The hot cathode furnishes a copious supply of electrons continuously. These electrons add greatly to the conduction of current (electron flow) and reduce the voltage drop required for producing simple ionization. The conduction of current is non-self-sustaining and the characteristics differ greatly from conduction in vacuum or self-sustaining conduction in gases. The characteristics of the hot-cathode gaseous tube will be covered in the chapter which follows.

Gaseous Lighting Units. Many lighting units of both novel and useful design operate on the principle explained in the preceding articles. One of the first of these is the familiar Geissler tube which usually consists of a long glass tube bent into irregular shape and containing a number of sections, each having a different kind of gas under low pressure. The sections are connected by conductors and the whole tube is placed in operation by the application of a high-voltage alternating current to the terminals. Very striking color effects are obtained when the tube is placed in operation and viewed in the dark.

The neon light consists of a long glass tube about ½ inch in diameter and filled with neon gas under low pressure. It is a very efficient source of light, giving a reddish-yellow color. The low cost and high efficiency of this light source has resulted in its popularity for sign lighting all over the world. Argon gas and sometimes helium and mixtures of gases are used to give different colors in sign lighting.

Glow lamps consist of bulbs filled with an inert gas such as neon under low pressure and containing two electrodes. These lamps use

the principle of ionization and the accompanying production of light arising from negative glow for illumination or for voltage regulation. One form of glow lamp is placed in an incandescent-type lamp bulb having a high resistance in its base. The resistance limits the magnitude of the current of gaseous conduction. These lamps consume from $\frac{1}{25}$ to 5 watts and are used for night lamps, signal lamps, and Christmas-tree lamps. Several lamps of this type are illustrated in Fig. 8.

The Cooper-Hewitt mercury-vapor lamp consists of a long tube containing a small quantity of mercury. Ionization of the mercury is

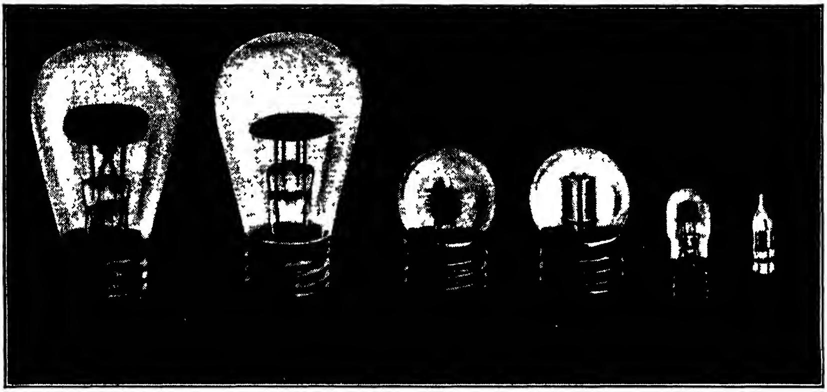


FIG. 8. Neon glow lamps. (Courtesy General Electric Company.)

started by tilting the tube, by applying a high inductive voltage, or by the use of thermionic emission from filaments. The color of the light, bluish yellow and devoid of red rays, causes human faces to have a deathly pallor. The specific advantages of this light are that it casts but little shadow, covers a narrow band in the visible light spectrum, and gives excellent visual acuity. Formerly the Cooper-Hewitt unit had fairly extensive application for drafting rooms, machine shops, and industrial processes where detection of color was not essential but where fine definition was important.

Today this low-pressure type of mercury-vapor lamp is being replaced by the fluorescent lamp which provides a lower cost and higher efficiency. Moderate- and high-pressure mercury-vapor lamps are concentrated and efficient sources of light that have many special applications.

The fluorescent lamp consists of a long glass tube coated on its inner surface with a fluorescent material known as phosphor and containing a small amount of mercury vapor. Emitter electrodes at the ends of the

tube produce initial electrons for ionizing the mercury vapor and furnishing ultraviolet light. The ultraviolet light causes the phosphor to fluoresce and give forth visible light. The different phosphors available are capable of producing different colors of light, including white, at relatively high efficiencies compared to the heated-filament type of light source.

The gaseous conduction of mercury vapor has given rise to a number of important developments. Much of the electromagnetic radiation from the ionized mercury vapor is ultraviolet light. Ultraviolet rays have a high therapeutic and sterilizing value. Accordingly, one type of mercury-vapor lamp (called a sun lamp) is built to give healthful light for man and animals. A special form of this lamp (germicidal) is used for sterilization in refrigerators, in operating rooms, and in treatment of wounds.

The sodium-vapor lamp uses the vapor of sodium as the gaseous conducting medium. A small quantity of metallic sodium is contained in an inner chamber. It requires several minutes for the lamp to heat up, evaporate the sodium, and come to full normal brilliancy. The lamps give a yellow light at a very high luminous efficiency and are used for highway lighting.

Chapter XII

GASEOUS AND VAPOR ELECTRON TUBES

Gaseous versus Vacuum Tubes. Two criteria for distinguishing between vacuum and gaseous electron tubes are (1) the pressure of the gas and (2) the mean free path of the electrons. In a highly evacuated tube of 10^{-8}-millimeter pressure at 0 degrees C, there are more than ten billion molecules in each cubic centimeter. For the ordinary vacuum tube of 10^{-6}-millimeter pressure, the mean free path of the electron is of the order of 42 meters. Since the electrode spacing in vacuum tubes never exceeds 2 or 3 centimeters (except in cathode-ray tubes), it is obvious that electrons move freely from cathode to anode with rare collisions with molecules of gas. In contrast with this picture, in gaseous tubes the normal spacing distances of electrodes is much greater (up to several inches), whereas the mean free path for common gas pressures of 10^{-2} to 10^{-1} millimeter is of the order of a *millimeter* or less. Thus in gaseous tubes the electrons suffer many collisions in transit and are likely to produce an intense ionization of the gas.

Neutralization of Negative Space Charge. Positive ions formed in an electron tube tend to neutralize the negative space charge surrounding a hot cathode. If a small amount of an inert gas is admitted to a high-vacuum diode, the gas molecules will be ionized by collisions with the emitted electrons in transit to the anode. The electrons formed by collision will be attracted to the positive anode and, because of their small mass, will reach it quickly. The positive ions formed by collision will move to the negative cathode rather slowly because of their large mass. Since the positive ions move slowly, they will remain in the cathode-anode space for a long time relative to the time of transit of electrons. Accordingly, a large number of positive ions may exist at a given instant within the region of the negative space charge surrounding the hot cathode. A single positive ion close to an electron will *neutralize the charge* of that electron at that instant. Thus the field produced by this pair is zero (neutral space charge). In an infinitesimal fraction of a second the electron of the pair is whisked away by the electric field and the positive ion moves a short distance toward

the cathode. In its new position the positive ion will be close to a second electron and at that instant will serve to neutralize this electron. In this manner the slow mobility of the positive ion will permit it to neutralize many (perhaps hundreds) of electrons while passing through the negative space-charge region. It is readily conceivable that, if at every instant the region surrounding the cathode contains the same number and distribution of positive ions and electrons, the negative space charge will be completely neutralized. This ideal state may seldom exist, but any degree of neutralization will overcome partially the negative space charge and will increase the number of primary electrons that are attracted to the anode. In passing, it should be noted that, if the positive ions present close to the cathode exceed the number of electrons, they will act like a positive grid very close to the cathode and thereby will increase greatly the primary electron flow to the anode.

The conduction of electricity in the gaseous hot-cathode tube may be considered as consisting of the three following components:

1. The primary electron current existing in a high vacuum.
2. The gaseous conduction resulting from electron transfer by ions.*
3. *The increased flow of electrons of thermionic emission resulting from the neutralization of negative space charge.*

The third component is the largest in magnitude and of the greatest importance. Component number two is usually small in magnitude. The addition of a small amount of gas or vapor greatly increases the current rectified by a diode. It also reduces the voltage required between the cathode and anode. This action is illustrated in the curves in Fig. 1*a*, giving a comparison of the anode-current anode-voltage characteristics of hot-cathode vacuum and gaseous tubes. In the gaseous tube with negative space charge neutralized, the anode current rises abruptly after the ionizing potential is reached, giving a constant anode-cathode drop with load. Thus for the gaseous tube the rectified power output is increased and at the same time the power loss within the tube is reduced. Both these changes increase the efficiency of rectification of alternating current.

Potential Distribution in Hot-Cathode Gaseous Tube. The presence of positive ions in gaseous and vapor tubes has a marked effect upon the potential distribution between the cathode and anode. In a high-vacuum tube the negative space charge depresses the voltage near the cathode, whereas in the gaseous tube the positive ions tend to neutralize

* This component is frequently termed the gas current.

this negative space-charge effect. Furthermore, if sufficient gas or vapor is present, an arc form of discharge takes place. The positive ions form a plasma for the arc which extends close to the cathode giving the potential distribution as shown in Fig. 1*b*. The trend of this curve is similar to that of the glow-discharge tube but here the magnitude of the cathode drop from plasma to cathode is much lower. In the glow-discharge tube a high voltage is necessary to produce electrons from a cold cathode, whereas in the tube under consideration a copious supply of electrons is emitted thermionically. In the gaseous and

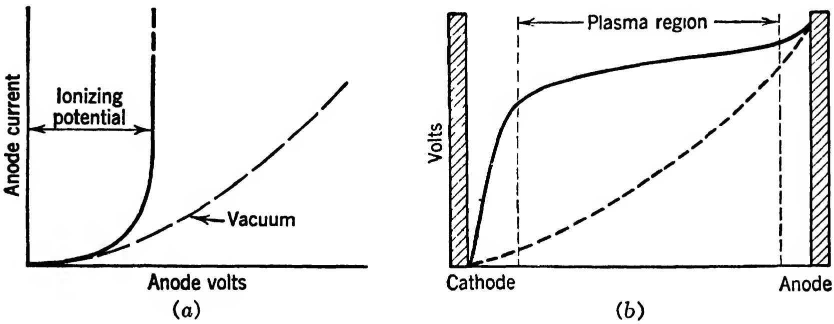


Fig. 1. Current and potential characteristics in a hot-cathode gaseous or vapor-rectifier tube.

vapor-arc tubes the total fall of potential from anode to cathode is of the order of 10 to 25 volts. The exact form of the potential distribution curve will vary with the gas or vapor and the pressure used and somewhat with the geometry of the tube. In general, there will be a rise of potential from the plasma of the arc to the anode for extracting the electrons out of the plasma. The potential distribution which would exist in vacuums because of negative space charge is indicated by the dotted line in Fig. 1*b*.

Cathode Sputtering. The positive ions produced in a gaseous tube bombard the hot cathode. When the ions hit the cathode they give up energy in two ways. First, the positive ions possess kinetic energy arising from the velocity acquired in the electric field. Secondly, the positive ion carries potential energy resulting from ionization. This latter energy is released when recombination takes place at the cathode. Both forms of energy given up when the positive ion hits the cathode may be transformed into heat, thus raising the temperature of the cathode and increasing the rate of thermionic emission. This process is utilized in the ionic-heated cathode and will be referred to later. If

the electric field is too high the bombardment of the positive ions may disintegrate the emitting surface. This action is known as *cathode sputtering* and the removed active material such as thorium or barium may land on a nearby electrode such as a grid and result in emission from that electrode. The loss of active material from the sputtered cathode will reduce the emission from the cathode and may make the cathode inoperative, thus ruining the tube for further service. Cathode sputtering results from (1) insufficient cathode emission (underheating), (2) too large cathode-anode current (overload), or (3) too high cathode-anode voltage drop. These conditions have a direct interrelation so that two of them occur simultaneously. Cathode sputtering can be prevented with a suitable resistor in the load circuit to limit the current output of the tube to the rated value and by having the cathode at normal emission temperature and the correct bulb temperature before anode voltage is applied.

While cathode sputtering is very harmful in an electron tube, the sputtering process may be useful in some manufacturing processes. Thus a thin coating of a metal may be placed on a plate or an electrode by sputtering this metal from a second electrode serving as a cathode.

Inverse Peak Voltage. The applications of any electron tube having unilateral conductivity depends largely upon the maximum inverse or reverse peak voltage that may be applied in the cathode-anode circuit without a reverse current flow. The term inverse voltage does not refer to the insulation or dielectric strength of the tube parts. In the high-vacuum tubes previously discussed it was pointed out that, with a cold anode, exceedingly high inverse voltage (anode negative and cathode positive) could be employed without any danger of a reverse current (called breakdown or arc-back). However, when gas is present in a tube conduction may take place from a negative cold plate, as was explained in Chapter XI, for the Townsend discharge, the glow discharge, or the arc discharge. Accordingly, in gaseous and vapor electron tubes the maximum inverse voltages must be held at much lower values.

The inverse voltage across a rectifying tube without a filter is the maximum value of the impressed a-c voltage. If a tube supplies a load through a filter, the maximum inverse voltage between the cathode and anode will be equal to the impressed a-c maximum voltage plus the d-c voltage at the input to the filter. This relation is illustrated for the half-wave rectifier and filter circuit shown in Fig. 2. The magnitude of the inverse voltage present will depend upon the filter and load circuit used and may approach a value twice the crest of the input.

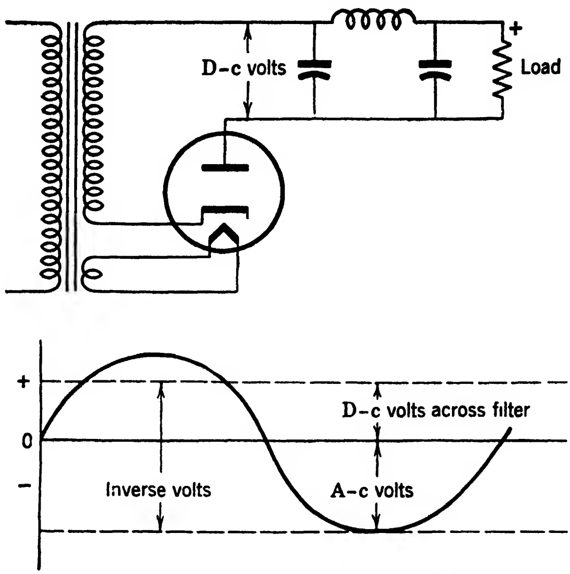


FIG. 2. Inverse voltage across a vacuum tube supplying a load through a filter.

Effect of Gas Pressure. Inert gases used in incandescent lamps reduce the evaporation of tungsten and permit higher operating temperatures and higher luminous efficiency for a given life. In a similar way, the use of inert gas in an electron tube reduces the evaporation from the cathode and permits the use of higher cathode temperature with corresponding increase of emission current. The higher the gas pressure, the greater the permissible thermionic emission.

The inverse peak voltage of a gaseous or vapor tube decreases rather rapidly with a rise in pressure. As the gas pressure rises above zero, the ionization by collision increases. The tendency to arc-back depends upon the number of positive ions present, and hence the permissible gas pressure will be determined by the inverse peak voltage required. In tubes employing mercury vapor the temperature of the tube becomes an important factor because the amount of mercury evaporated, and hence the vapor pressure, depends upon the coolest place in the tube where the vapor condenses. Thus in a mercury-vapor tube the operating temperature may become very critical as far as peak inverse voltage is concerned.

Gaseous Rectifier Diodes. The Tungar gaseous rectifier diode is filled with argon gas under a pressure of approximately 2 pounds. The argon gas furnishes the positive ions for neutralizing negative

space charge and the high gas pressure reduces evaporation and thorium loss from the cathode. The anode is a graphite disk with its lead brought in from the top of the tube. The cathode is a helical coil of tungsten or thoriated-tungsten wire connected through leads to a screw type of base, as illustrated in Fig. 3. The cathode filament is operated at a low voltage of 1.8 to 2.2 volts and with currents of 6 to 27 amperes, depending upon output current. The tube "fires" or conducts with a nearly instantaneous start. No harm results from impressing voltage on the anode before the cathode reaches operating temperature because the relatively high gas pressure limits the velocity (collisions) of the positive ions. The average d-c pick-up voltage for the tube is 11 to 13 volts and the average arc drop (cathode to

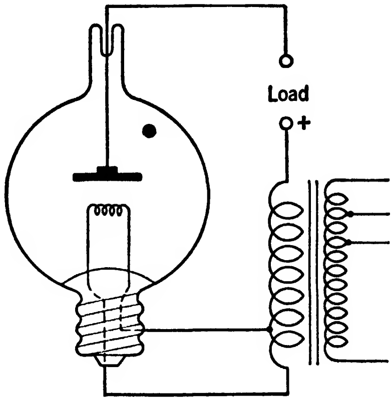


FIG. 3. Circuit for a half-wave Tungar rectifier.

RECTIFIER EL6B

(Half-wave)

TANTALUM ANODE AND XENON GAS

Cathode	
Filament voltage	2.5 volts
Filament current	21 ± 2 amps
Heating time	1 min
Anode	
Peak inverse voltage	920 volts
Arc drop (average)	9 volts
Starting voltage (average)	12 volts
D-c current (continuous)	6.4 amps
D-c current (10 seconds)	12.8 amps
Ambient temp. limits	−55 to 70 C

Note: The cathode should be heated before the load voltage is applied.

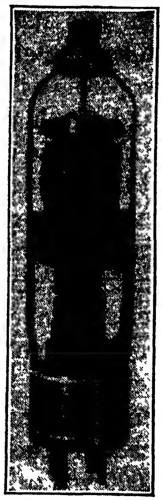


FIG. 4. Typical gaseous rectifier tube. (Courtesy Electrons Inc.)

anode) is only 7 volts. The maximum inverse peak voltage varies from 150 for small-capacity tubes to 300 for the 6-ampere capacity. This low inverse peak is necessary to prevent a glow discharge from

starting on the nonconducting half of the cycle of rectification.

A second type of gaseous diode (Fig. 4) uses pure xenon gas at a pressure of less than one millimeter of mercury. The reduced pressure permits a higher inverse peak voltage. Xenon gas has practically the same space charge neutralizing power as mercury, and its density is unaffected by the surrounding temperature. A lightweight tantalum anode permits thorough degassing in manufacture. Gaseous rectifier tubes have output ratings varying from 0.5 to 50 d-c amperes and are used for charging storage batteries, arc supply for motion-picture projectors, and other moderate capacity d-c loads.

Mercury-Vapor Rectifier Diode. The mercury-vapor rectifier diode uses a hot cathode and mercury vapor under a *low pressure*. It should not be confused with the pool type of tube which uses a pool of cold metallic mercury for a cathode. A small quantity of mercury is inserted in a hot-cathode evacuated tube. A part of or all the mercury vaporizes and the vapor atoms ionize and serve as the conducting medium in the tube. The conduction is of the arc type. In comparison with the vacuum diode the mercury-vapor tube carries a much larger current, has a neutralized negative space charge, and has a nearly constant cathode-anode voltage drop (within the range of 10 to 20 volts) which is nearly independent of current. The absence of negative space charge and its accompanying losses allows a larger electrode spacing and smaller-size electrodes for a given current-carrying capacity. It also permits the use of an electron-emitting cathode of higher efficiency and much larger current-carrying capacity. The anodes consist of disks made of metal or graphite which are mounted on leads brought out at the top of the tube. The hot cathodes are of the oxide-coated type and are generally placed within a hollow metal cylinder to reduce radiation losses. Various configurations have been used in the construction of these cathodes as illustrated in Fig. 5.

Glass-enclosing tubes are used for mercury-vapor units having small current-carrying capacities. Metal-enclosing tubes are used for the larger capacities.

The mercury-vapor rectifier is superior to the gaseous rectifier in that it will withstand much higher inverse peak voltages of the order of 5,000 to 10,000 volts. However, the mercury-vapor diode has the disadvantage that voltage should not be applied to the anode until the cathode and the entire tube reach normal operating temperature. Such application of voltage when the cathode is cold or partially heated will result in the destruction of the cathode-emitting surface by positive ion bombardment. Special care in the installation and operation of mer-

cury-vapor tubes is necessary to insure long life and satisfactory operation. The user should observe carefully the manufacturer's instructions for operation in an upright position, long warm-up period when

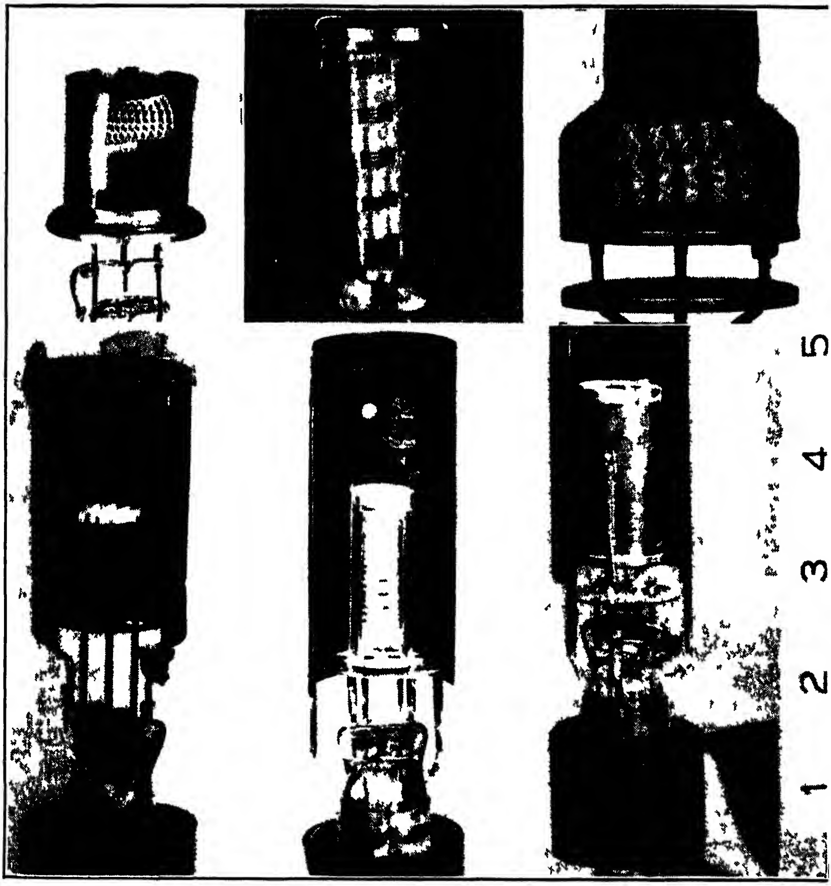


FIG. 5. Assemblies of oxide-coated cathodes used in gaseous and vapor tubes. (Courtesy General Electric Company.)

first installed, suitable surrounding temperature, protection from water drops, acid fumes, and so forth.

The accepted name for the mercury-vapor diode in the power industry is the *phanotron*. These tubes have average current output ratings of from 0.25 to 30 amperes. A tube of low rating is shown in Fig. 6 and one of high current in a metal enclosure in Fig. 7. Phanotrons are used for supplying direct current for applications where medium values of voltage and current are required.

Some diodes termed Tungars contain mercury vapor and operate at a higher pressure comparable to those containing inert gas. These mercury-vapor tubes possess characteristics similar to those containing gas. Similarly some diodes containing inert gas operate at pressures comparable to phanotrons and possess similar characteristics. Other diodes contain both argon gas and mercury vapor. These tubes possess a low pick-up voltage (about 4 volts) and a low inverse peak (about 55 to 75 volts).

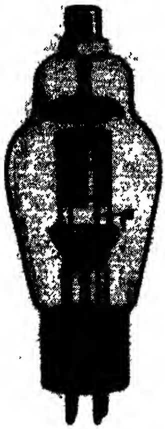
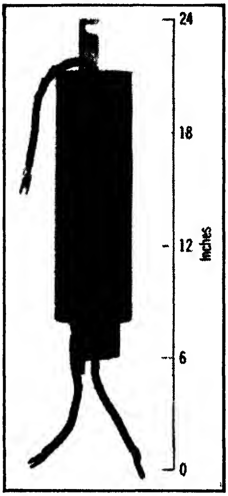


FIG. 6. Mercury-vapor rectifier tube (average anode current 0.25 ampere). (Courtesy Radio Corporation of America.)

• **Thyratrons.** The thyratron is a triode containing inert gas or mercury vapor under low pressure. It is a phanotron plus a control grid. The admission of gas or vapor into a three-electrode vacuum tube greatly changes the characteristics and operation of the device. The presence of gas renders the tube useless as a grid-controlled amplifier but makes it valuable as a grid-controlled arc rectifier. In order to perform this new function the grid should surround the cathode or the anode so as to screen the one from the other. One early type of construction of the thyratron is shown schematically in Fig. 8.

The theory of action of the thyratron can be pictured by considering its behavior in a d-c circuit. Let the filamentary cathode of Fig. 8 be heated to normal temperature for emission and let switches S_p and S_c be open. Close S_p, placing a high positive potential on the anode with the grid free, and nothing will happen if the grid has a fine mesh. This follows because the free grid is made negative by the electrons from the cathode landing on it. This negative grid repels the electrons emitted by the cathode so that few get past the meshes of the grid to start ionization. Next, if switch S_c is connected to the positive side of the C battery, electrons pass through the grid readily, ionization starts, an arc develops, and the tube conducts an electric current which is limited by the magnitude of the load resistance. After the ionization has started, the opening of switch S_c, causing a free grid to exist, has no effect on the cathode-anode arc. Likewise, if the switch S_c is connected to the negative side of the C battery, the resulting negative charge on the grid has no appreciable effect on the arc. Thus the grid has the power to start the arc but no power to control its magnitude or to stop it after it has started. The reason for this unexpected action lies in

the presence of the positive ions in the arc. These ions, which fill the cathode-anode space, constitute the plasma of the arc. When the grid is made negative, it attracts some of the positive ions from the plasma



GENERAL CHARACTERISTICS

Number of Electrodes		[illegible]
Electrical		
Cathode-filamentary		
Filament voltage		2.5 volts
Filament current, approx		100 amp
Heating time, typical		2 min
Optimum phase of filament voltage with respect to anode voltage		90 degrees
Peak voltage drop, typical		9 volts
Rise of tube temperature above ambient		
Without forced-air circulation		
For average anode current	0	20 amp
Degrees rise, condensed mercury temperature	30	35
Degrees rise, temperature of side of tube	100	150
Maximum ratings		
Maximum peak inverse anode voltage		
20° to 60° C condensed mercury		1500 volts
20° to 70° C condensed mercury		800 volts
Maximum anode current		
Instantaneous		75 amp
Average		20 amp
Surge, for design only		750 amp
Maximum time of averaging current		30 sec
Temperature limits, condensed mercury +40° to +60° C		

FIG. 7. Phanotron FG-166. (Courtesy General Electric Company.)

and these positive ions fly to and cluster around the grid like bees coming to a hive (Fig. 9). Each positive ion seizes an electron and becomes a neutral atom. However, the process is continuous and the

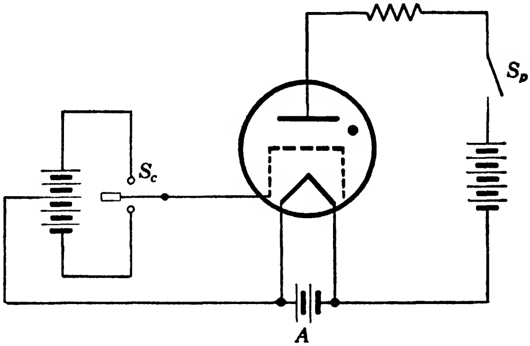


FIG. 8. Thyratron in a d-c circuit.

cluster or sheath of positive ions surrounding the grid creates a positive field (space charge) as far as the surrounding region is concerned. If the grid is made more negative, a thicker sheath of positive ions

surrounds it but without any effect on the cathode-anode arc, if the meshes of the grid are some distance apart. However, if the grid wires are very close together (almost touching), the sheath of positive ions around the grid will serve to limit the cathode-anode arc. Thus, theoretically, a closely spaced grid plus a high negative grid potential can be used to limit and even to stop an arc. Such means are not employed in practice. The simple and usual method for interrupting the arc when using direct current is to break the cathode-anode circuit by opening the switch S_p. The arc will be interrupted also if the anode voltage is reduced below the minimum required to maintain ionization.

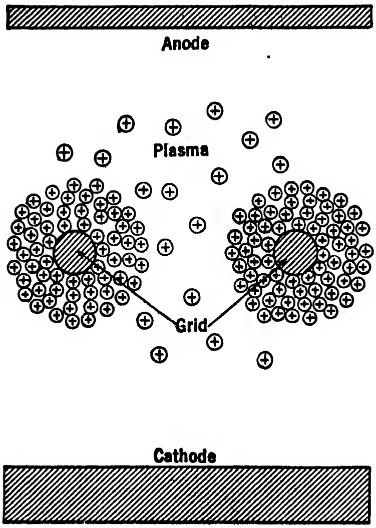


FIG. 9. Grid action in a thyratron.

The potential required on the grid of a thyratron to permit the rectifying arc to start depends on several factors. These factors are the size of the openings of the grid, the gas or vapor pressure within the tube, the general geometry of the tube, the grid current and circuit resistance, and the potential applied to the anode. The voltage conditions for starting the arc depend to a large degree on the structure of the grid. Thus with a fine-mesh grid (small hole or holes) a positive potential must be applied to the grid to start gaseous conduction, whereas with a coarse grid (large hole or holes) the arc may start with a negative potential on the grid. For the range of gas pressures used on thyratrons the low pressures require a higher potential gradient to start the arc. The positive potential on the anode determines the electric field within the tube and hence the "pull" upon the electrons. Thus a high initial positive voltage on the anode will require a more negative potential on the grid to prevent the formation of an arc. For a given tube, there is a certain ratio of anode volts to grid volts at which the tube will "fire." This ratio is called the grid-control ratio. It can be expressed as

$$\rho = -\left(\frac{e_a}{e_c}\right) \qquad i_a = 0$$

and it bears a resemblance to the amplifying factor mu of the vacuum triode. The grid-control characteristics of thyratrons from which the grid-control ratio can be obtained are illustrated in Fig. 10. The grid-control ratio for a tube with the characteristic A is cy/dy, and this ratio will be nearly constant along the linear portion of the curve.

Thyratrons may be classified by the sign of the grid-control voltage that permits the starting of the tube. A negative-control tube is illustrated by curve A of Fig. 10 where a negative potential must be applied to the grid to prevent starting for all values of anode potential within the range of operation. For a similar reason, curve C represents a tube having a positive control. Curve B of Fig. 10 represents a tube of the intermediate-control class. This tube requires a positive grid for starting for anode voltages up to point X and a negative grid for higher values. This class of tube has a short deionization time and is used for inverter circuits.

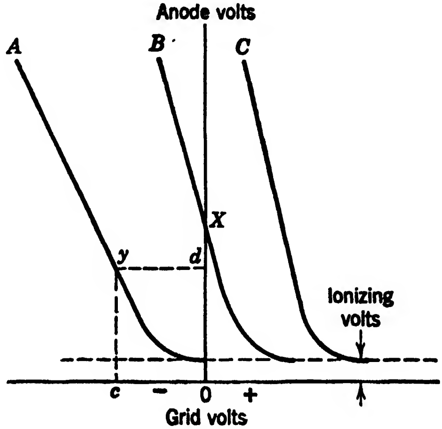


FIG. 10. Anode grid-voltage characteristic of thyratrons.

The thyratron is a "natural" for use on a-c circuits. If the grid is given a potential sufficiently positive so that the tube will conduct for all values of anode potential (above ionizing potential), the tube will rectify all positive loops of a-c potential exactly like a gaseous diode. However, if the potential of the grid of the thyratron is controlled in a suitable manner the thyratron may be made to conduct current or to stop current at will. This follows because the a-c voltage passes through zero twice during each cycle and is negative for half of each cycle. Thus if the grid of the thyratron is brought to the critical value, as determined by the grid-control ratio, the arc is struck each time the anode voltage becomes positive, but, if the grid voltage is lowered (made more negative), the current will stop at the end of the positive anode-voltage loop. Then when the anode voltage becomes positive on the next and succeeding cycles, the tube does not "fire." This simple grid control for starting and stopping current rectification

in the anode circuit has given rise to the term "trigger tube" for the thyratron. The operation of the thyratron on alternating current is illustrated in Fig. 11. The grid voltage necessary to stop conduction for any given anode potential is given by the grid-control characteristic on the left. The necessary restraining values can be referred to the impressed a-c voltage cycle as shown by the graphical construction of Fig. 11. The curve e_c shows the necessary values of grid po-

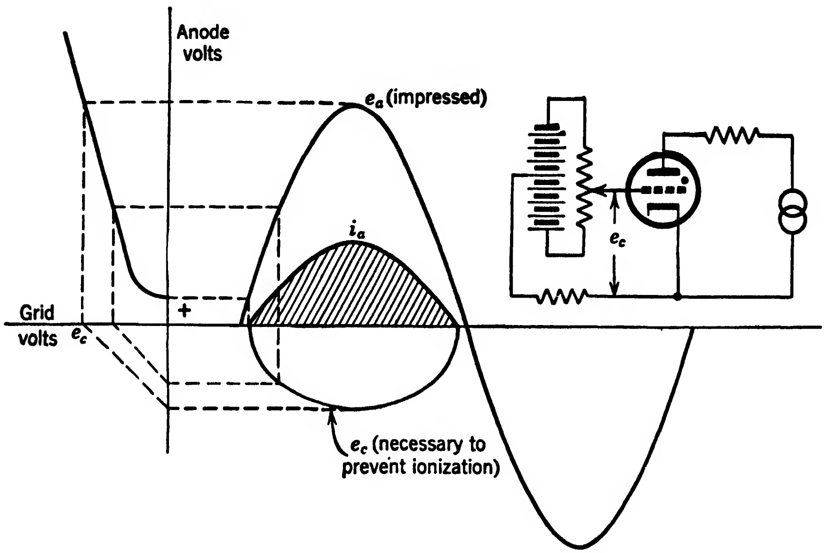


FIG. 11. Action of a thyratron on alternating current.

tential to prevent conduction for corresponding values of positive anode potentials. For a complete "stop" current, the grid must have a negative value greater than the maximum of the e_c curve.

The e_c curve of Fig. 11 suggests that the grid control of the thyratron may be used to control the magnitude of the rectified current as well as for "on" and "off" operation. Thus if the grid is held sufficiently negative until a certain part of each positive anode-voltage loop is reached, the tube will fire and conduction will occur for the remainder of each positive half of the cycle. The grid may be made to exercise this form of control in two ways, known as *amplitude control* and *phase-shift control.* Amplitude control is brought about by applying a negative bias to the grid as illustrated in Fig. 12. For the bias shown under *a* the thyratron conducts current beginning at point X and the area of the rectified current loop is reduced slightly. An increase in the

bias to the value shown under b will permit the tube to fire at point Y so that the magnitude of the rectified current has been reduced to one-half the normal value. Values of negative bias between those shown will give other magnitudes of current. Obviously, this is a very simple method of controlling current values but it has some limitations and disadvantages. First, the maximum reduction in magnitude is one-half value (one-fourth cycle). This maximum reduction is critical

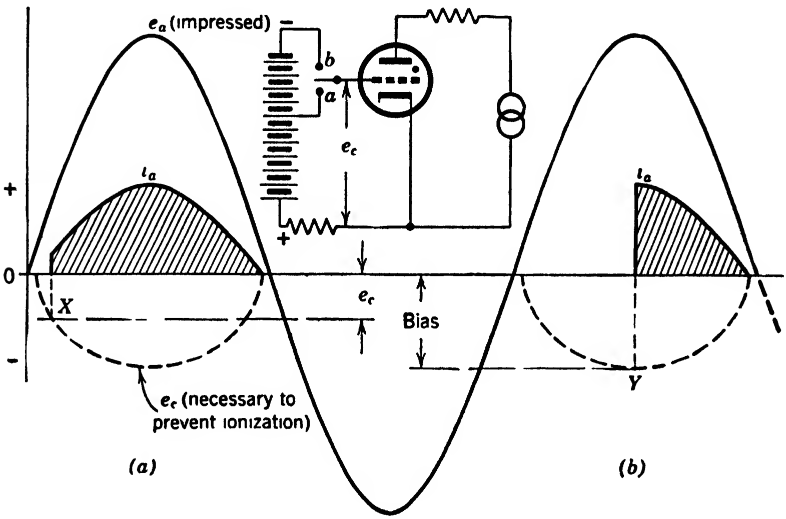


FIG. 12. Amplitude grid control of a thyratron.

because the firing point depends on both the grid and anode potentials and a temporary decrease in anode voltage might cause the tube to skip some cycles. For like reasons changes in anode potential, thermionic emission from cathode, and changes in temperature will vary the firing point, and hence the control of the current magnitude for a given grid-bias setting may not be sufficiently exact for some applications.

A more exact timing and firing of the thyratron can be secured by phase-shift control. In this method an a-c voltage derived from the anode supply line is applied to the grid. The phase of the grid voltage is shifted with respect to the anode voltage by suitable means, giving results as shown in Fig. 13. If the grid voltage e_g is "in phase" or ahead of the anode voltage as shown in part a, the tube fires as soon as the positive loop of the anode potential reaches the ionizing voltage and the complete loop is rectified. A delayed shift of the grid voltage e_g as shown in part b will cause the rectification to begin at the end of one-

third of the positive loops. A further shift in the grid voltage will delay the firing to a point near the end of the positive loop, as illustrated in part *c*. A shift of the grid potential from "in phase" to a lag of 180 degrees will permit a control from "full" to zero magnitude of current. Any change of anode voltage will be reflected in a change in grid

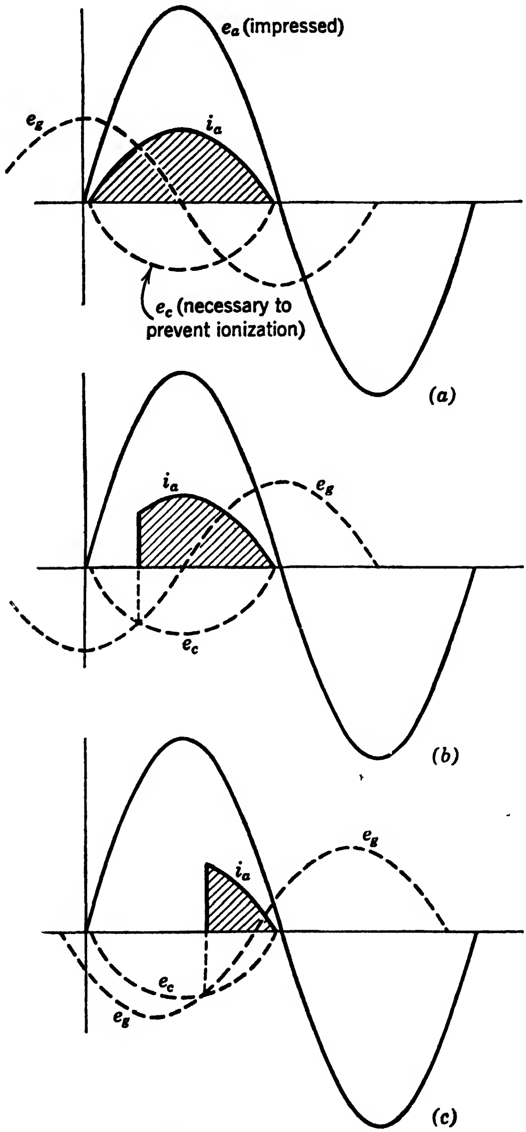


FIG. 13. Phase-shift grid control in a thyratron.

voltage and in the value of the e_0 curve so that the period in the cycle where the firing occurs will remain nearly constant. For some applications it is advantageous to superimpose the controllable a-c shift component of the grid voltage upon a normal grid bias, as shown in Fig. 14. Here the grid bias is made sufficiently negative to prevent the tube from firing if the a-c grid potential is not applied. Since the grid is always negative, the power required for the grid input is reduced to a very low value. Another useful combination of a-c and d-c grid voltages for firing thyratrons is to use an a-c voltage component with a fixed phase shift (such as 90 degrees) and then control the firing angle by a variation of the d-c bias (vary the grid bias in Fig. 14).

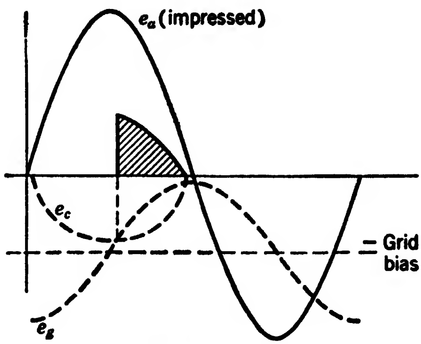


FIG. 14. Phase-shift control superimposed upon a normal grid bias in a thyratron.

The phase shift for the grid circuits of thyratrons and other electronic devices may be produced by mechanical devices or by simple circuits, one of which is illustrated in Fig. 15. The theory of this phase-shifting circuit is illustrated in the vector diagram of Fig. 15. The a-c voltage impressed across the cathode-anode circuit is in phase

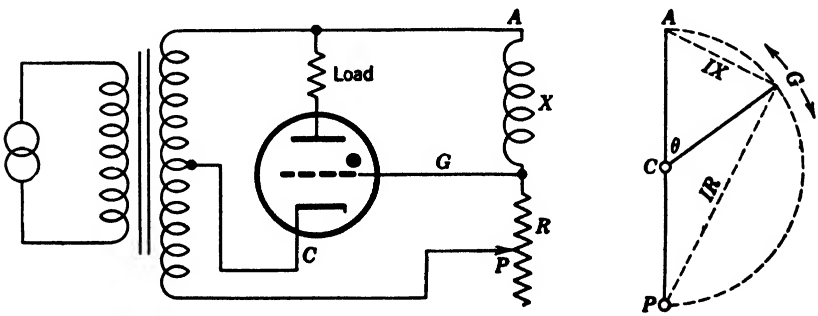


FIG. 15. Simple phase-shift circuit and vector diagram.

with that across points A and P. With a very high value of resistance R, the point G on the vector diagram will show voltage CG (cathode-to-grid) nearly in phase with that across the cathode-anode circuit.

Now as R is reduced, the magnitude and phase of AP remains unchanged but the position of G will swing clockwise on the arc of a circle, thus throwing the cathode-grid voltage out of phase with the cathode-anode voltage and making possible a 180-degree phase shift. For all positions of shift the magnitude of CG remains constant. A condenser may be substituted for the inductance of Fig. 15 to give another simple phase-shifting circuit. In making this change, the relative positions of R and C must be interchanged in order to retain a lagging phase shift.

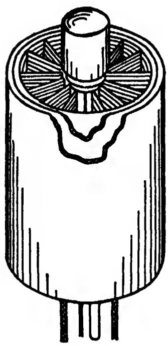


FIG. 16. Heater type of cathode used on thyratrons.

Some small-capacity thyratrons have cathodes of the filament type as in vacuum tubes. For the larger thyratrons there are two objections to this type of cathode. In the first place, the filament has a large heat loss and is very inefficient. In the second place, the voltage drop between the ends of filament reduces the available range of the cathode-anode potential drop. The neutralization of the negative space charge by the presence of positive ions permits a greater distance between cathode, grid, and anode and also makes possible the confinement of the cathode in a small space or in a heat-insulated oven. Several examples of cathodes used in thyratrons are illustrated in Fig. 5 and an additional shield and heater type is shown in Fig. 16.

The grids of thyratrons are placed farther from the cathode than in vacuum tubes to make the grid control more effective and to keep the grid at a lower temperature so that grid emission will not render the control ineffective. Several types of grids have been used in thyratrons. Grids in early tubes were hollow cylinders with one end closed. The walls of the cylinders consisted of (1) a wire mesh, or (2) sheet metals with small holes for the passage of electrons and ions (parts *a* and *b*, Fig. 17). The basic type for the grids used today consists of a hollow metal cylinder with open ends but closed by a baffle near the center. The anode and cathode are placed within the cylinder, the former above the baffle and the latter below. Holes for the passage of electrons and ions may be placed in the sides of the grid cylinder or in the baffle plate only (parts *c* and *d*, Fig. 17). For thyratrons in metal enclosures, the enclosure itself constitutes the grid structure.

Inert gas such as argon, xenon, or mercury vapor is used in thyratrons at pressures varying from 1 to 50 microns (a micron is 10^{-6} meter of mercury). It is essential that the gas be pure. In normal operation

of a mercury-vapor arc there are 10^{+10} to 10^{+12} ions present per cubic centimeter. The gas in the tube aids in carrying the heat away from the anode.

Thyratrons are built in many sizes and ratings. Those of low capacity are contained in glass tubes and use inert gas. Those of large capacity are contained in metal tubes and use mercury vapor. It is possible to build thyratrons with current ratings as high as 100 amperes and with voltages up to 100,000. However, for large current capacities it has been found preferable to use other types of tubes to carry the load current and to utilize the thyratron for controlling the second tube. A miniature thyratron for control is illustrated in Fig. 18. The construction, rating, and characteristics of two typical larger thyratrons are given in Figs. 19 and 20.

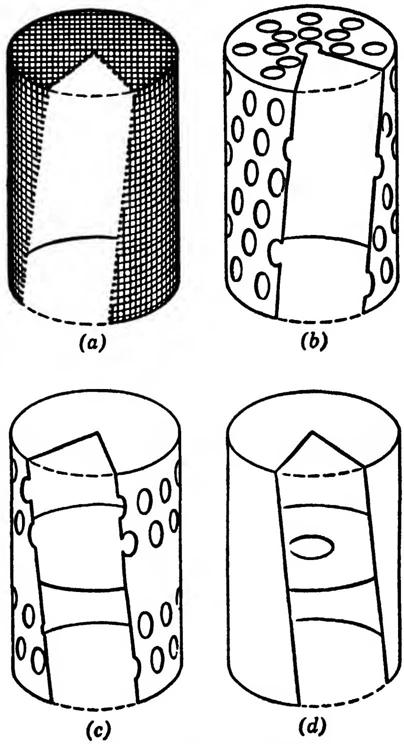


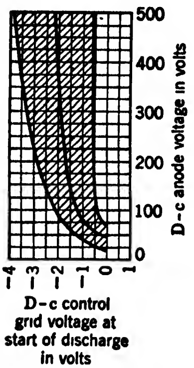
FIG. 17. Grid structures for thyratrons.

The efficiencies of thyratrons in voltage ratings of 500 and up are of the order of 97 per cent, which is excellent compared with other electrical machines and devices.

The precautions previously stated regarding the operation of mercury-vapor diodes should be observed in the use of thyratrons. The cathode should be brought up to normal operating temperature before voltage is applied to the anode to prevent cathode sputtering. The output current must not exceed the emission current. The cathode-anode voltage must be kept below the disintegrating potential for the cathode (about 28 volts). Both the current output and the cathode-anode voltage should be limited by a suitable load impedance.

It is usually necessary to place a resistor in the grid circuit of three-electrode thyratrons. There are a number of reasons for this practice. Since the grid lies within the path of the electron and ion flow, a fairly

large current may flow in the grid circuit for the tube that requires a positive grid or for any tube during the ionization period. When an a-c voltage is impressed across the grid, a grid current flows because of a rather large interelectrode capacity between the cathode and grid. Grid current in the thyratron produces some heating, changes the input impedance to the grid, and is likely to interfere with the operation



GENERAL CHARACTERISTICS

Number of electrodes	
Electrical Design	
Cathode—indirectly heated type	
Voltage	6.3 volts
Current, approx	0.15 amp
Heating time, typical	10 sec
Peak voltage drop, typical	11 volts
Average anode-to-control-grid capacitance	0.1 $\mu\mu$f
Mechanical Design	
Net weight, approx	1/8 oz
Operating position—any	
Maximum overall length	1 1/2 in.
Maximum overall diameter	11/16 in.
Maximum ratings	
Maximum peak anode voltage	
Inverse	500 volts
Forward	500 volts
Maximum anode current	
Instantaneous	100 ma
Average	20 ma
Maximum time of averaging anode current	15 sec
Ambient temperature limits	$-40°$ to $+80°$ C

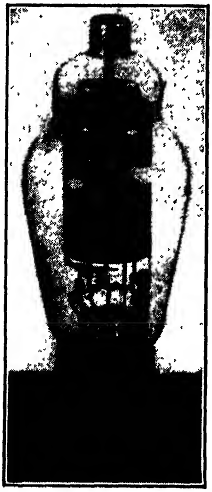
Fig. 18. Miniature thyratron GL-5663. (Courtesy General Electric Company.)

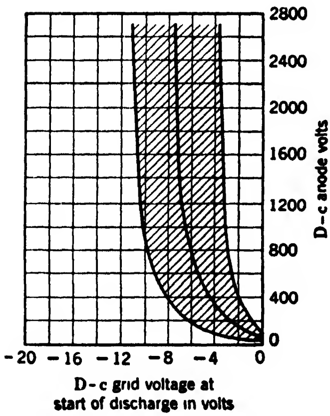
of the grid-control circuit. Grid series resistors have magnitudes in the range of $\frac{1}{4}$ to 2 megohms.

The thyratron with three electrodes has certain limitations somewhat similar to those of the vacuum triode which can be overcome through the use of a fourth electrode called a shield grid corresponding to the screen grid in the vacuum tetrode. The shield grid is usually a hollow metal cylinder containing two circular disk baffles as shown in Fig. 21. Each baffle contains a hole for the passage of electrons and ions. The control grid is a small ring placed in the conduction path midway between the baffles of the shield grid. The shield grid is maintained at a constant d-c voltage which tends to stabilize operation and to give the desired grid-control ratio.

The shield-grid construction serves to minimize grid current from a variety of causes and to permit satisfactory operation of the tube in a

high-impedance grid circuit. These advantages are accomplished in three ways. First the shield grid reduces both the anode-to-control-grid and the control-grid-to-cathode interelectrode capacity. The former is important in thyratron circuits since the steep wave-front





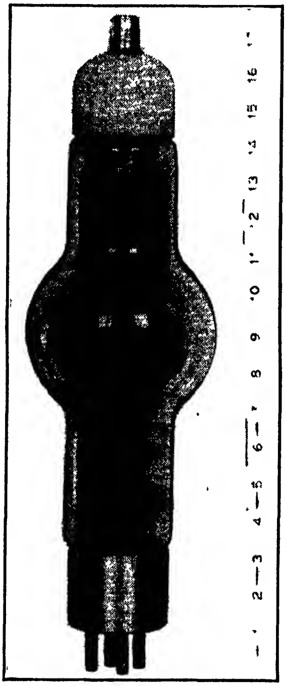
GENERAL CHARACTERISTICS

Number of Electrodes			3
Electrical Design			
Cathode—filamentary			
Filament voltage			2.5 volts
Filament current, approx			5.0 amp
Heating time, typical			5 sec
Peak voltage drop, typical			16 volts
Approximate control characteristics			
Anode voltage	40	100	1000 volts
Control-grid voltage	0	−2.25	−6.5 volts
Anode-to-control-grid capacitance			4.4 μμf
Deionization time, approx			1000 μsec
Ionization time, approx			10 μsec
Maximum ratings			
Maximum peak anode voltage			
Inverse			5000 volts
Forward			2500 volts
Maximum negative grid voltage			
Before conduction			500 volts
During conduction			10 volts
Maximum anode current			
Instantaneous, 25 cycles and above			2.0 amp
Instantaneous, below 25 cycles			1.0 amp
Average			0.5 amp
Surge, for design only			40 amp
Duration of surge current			0.1 sec
Maximum grid current			
Instantaneous			0.25 amp
Average			0.05 amp
Maximum time of averaging current			15 sec
Temperature limits, ambient			+40° to +80° C

FIG. 19. Thyratron FG-17. (Courtesy General Electric Company.)

transients frequently encountered in these circuits may be transmitted to the control grid through this capacity causing premature firing. The reduction in the cathode-to-control-grid capacity is of lesser importance though it does reduce the grid current input where a-c potentials are applied. The second advantage of the shield grid is its action in

shielding the control grid from contamination and temperature. The shielded position reduces the amount of material evaporated or sputtered from the cathode and anode which may become deposited on the control grid. Also the shielded position reduces the radiated heat



GENERAL CHARACTERISTICS

Number of Electrodes				3
Electrical Design				
Cathode—indirectly heated type				
Heater voltage				5.0 volts
Heater current, approx				20 amp
Heating time, typical				5 min
Tube voltage drop, typical				16 volts
Approximate starting characteristics				
Anode voltage	1,000	10,000	250 approx	volts
Control-grid voltage	−1.5	−4.5	0	volts
Deionization time, approx				100 μsec
Ionization time, approx				20 μsec
Maximum Ratings				
Maximum peak anode voltage				
Inverse				10,000 volts
Forward				10,000 volts
Maximum negative control-grid voltage				
Before conduction				1,000 volts
During conduction				15 volts
Maximum anode current				
Instantaneous, 25 cycles and above				75 amp
Instantaneous, below 25 cycles				25 amp
Average				12.5 amp
Maximum control-grid current				
Instantaneous				5.0 amp
Average				1.0 amp
Maximum time of averaging current				30 sec
Temperature limits, condensed mercury				40° to 65° C

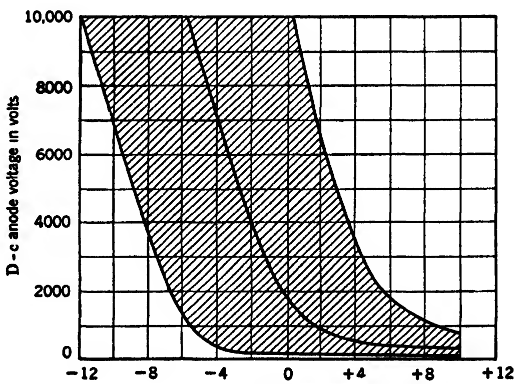


FIG. 20. Thyratron FG-41. (Courtesy General Electric Company.)

and lowers the temperature of the control grid which, in turn, reduces grid emission. A third advantage of the shield grid is to permit the use of a small control grid which reduces both its emission current and any current arising from the interelectrode capacity. It should be noted that the shield-grid current does not pass through the control-grid circuit.

GENERAL CHARACTERISTICS

Electrical Design		
Number of electrodes		4
Cathode type—indirectly heated		
Voltage	5.5	5.0 volts
Current, approx	11.0	10.0 amp
Heating time, typical		5 min
Peak voltage drop, typical		16 volts
Approximate starting characteristics		
Anode voltage	100	2000 volts
Shield-grid voltage	0	0 volts
Control-grid voltage	+1.0	−14.0 volts
Anode to control-grid capacitance		0.07 μμf
Deionization time, approx		1000 μsec
Ionization time, approx		10 μsec
Maximum Ratings		
Maximum peak anode voltage		
Inverse	750	2000 volts
Forward	750	2000 volts
Maximum negative control-grid voltage		
Before conduction		1000 volts
During conduction		10 volts
Maximum negative shield-grid voltage		
Before conduction		300 volts
During conduction		5 volts
Maximum anode current		
Instantaneous, 25 cycles and above	77	40 amp
Instantaneous, below 25 cycles		13.0 amp
Average	2.5	6.4 amp
Surge, for design only		400 amp
Duration of surge current		0.1 sec
Maximum control-grid current		
Instantaneous		1.0 amp
Average		0.25 amp
Maximum shield-grid current		
Instantaneous		2.0 amp
Average		0.50 amp
Maximum time of averaging current		15 sec
Temperature limits, condensed mercury	+30° to +95°	+40° to +80° C
Recommended temperature, condensed mercury		+40° C

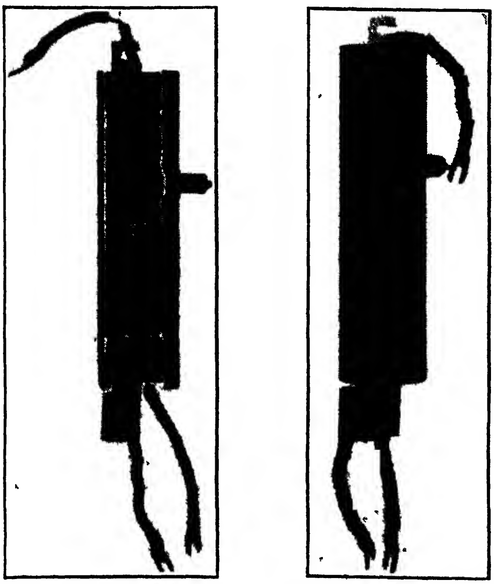


FIG. 21. Shield-grid thyratron FG-172. (Courtesy General Electric Company.)

The thyratron is one of the most useful control devices invented in the twentieth century. Its applications are too numerous to mention. In one application it serves as a relay. A small change of potential on the grid starts or stops a rectified current which is performing some useful function. A constant temperature is maintained in an electric furnace or oven by a thermostat which controls the phase shift or the

Peak anode voltage, max	8.0 kv
Peak anode current, max	90 amp
Peak inverse anode voltage, max	6.0 kv
Average anode current, max	100 ma d-c
Pulse duration (measured at ½ amplitude), max	6.0 μsec
Pulse repetition frequency, max	4000 pps

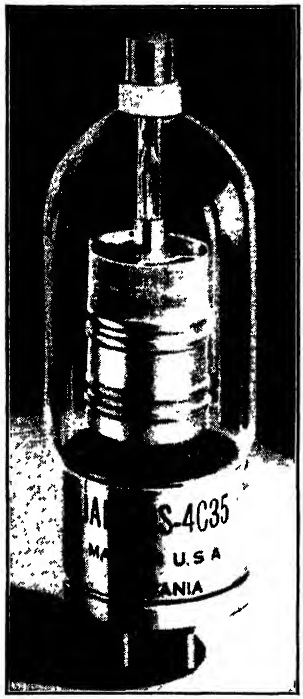


FIG. 22. Hydrogen thyratron. (Courtesy Sylvania Electric Products, Inc.)

on-and-off potential on the grid of the thyratron which furnishes rectified current to the device. Lighting circuits may be dimmed or lighted slowly or turned on and off by thyratrons. Here the thyratron varies the flux (saturation) in an iron-core reactor which is in series with the lighting circuit. The armature of a d-c motor may be supplied with direct current rectified by a thyratron, and the starting, stopping, and speed control may be governed by the voltage phase shift on the grid of the tube. The grid circuits of thyratrons may be energized by light falling upon a photocell.

An important factor in the use of the thyratron as a trigger tube is the time required for ionization and deionization. Ionization time for

commercial thyratrons varies from 10 to 20 microseconds and the deionization time from 100 to 1000 microseconds. While these periods appear to be short, they are long enough to limit the use of the thyratron for many high-frequency applications. It is possible to reduce the ionization time for mercury-vapor thyratrons by using peaked grid voltages of relatively large magnitude for firing the tube. Where very short ionization and deionization time is important the hydrogen thyratron may be used.

The hydrogen thyratron is a hot-cathode grid-controlled gas rectifier tube developed during World War II for pulsing service at high repetition frequencies, high peak currents, and high voltages. Its outstanding feature is the short deionization time required to convert the gaseous ions (hydrogen) to neutral molecules when the tube is shut off. This action results from the relatively small mass of the hydrogen molecule. Another advantage of the tube is that it may be operated over a wide range of ambient temperatures without significant change in electrical characteristics. The hydrogen thyratron has been applied in high-frequency and pulsing service but may find other applications where its characteristics are useful. A commercial hydrogen thyratron and its rating are shown in Fig. 22.

Load Current Values in Thyratron Circuits. In the application of thyratrons it may be desirable to know the average or effective values of the current in the load circuit for different values of the firing angle. If the assumption is made that the impressed a-c voltage has a sine wave and the load is purely resistive, the methods of the calculus may be employed. The average current (direct current) for any given firing angle will be proportional to the area under the rectified current wave divided by the base for one cycle. The area under the curve is

$$\text{Area} = \int_{\theta_1}^{\theta_2} I_m \sin \theta \, d\theta$$

and the average value in terms of I_m is

$$\text{Average value} = \frac{\text{area}}{2\pi}$$

Obviously, suitable mathematical manipulations must be applied to care for different firing angles and to care for partial half-wave or full-wave rectification (using two thyratrons).

The calculation of effective or rms values requires that the y ordinates of the sine wave be squared in determining the effective area,

and then the square root of the average must be obtained. The process involves the following approximate steps.

$$\text{Effective area} = \int_{\theta_1}^{\theta_2} I_m^2 \sin^2 \theta \, d\theta = I_m^2 \left[\frac{\theta}{2} - \frac{\sin 2\theta}{4}\right]_{\theta_1}^{\theta_2}$$

$$\text{Effective current} = I_m \sqrt{\frac{\text{effective area}}{2\pi}}$$

Calculations based on the preceding suggestions give the results shown in Table 1 for various firing angles and for the use of half- and

TABLE 1

Firing Angle (delay from zero time)	One Thyratron (half-wave)			Two Thyratrons (full-wave)		
	Average value in terms of I_m	Effective (rms) in terms of I_m	Ratio $\frac{\text{average}}{\text{effective}}$	Average value in terms of I_m	Effective (rms) in terms of I_m	Ratio $\frac{\text{average}}{\text{effective}}$
0	0.318	0.500	0.635	0.636	0.707	0.900
15	.313	.498	.630	.626	.705	.890
30	.297	.489	.610	.594	.695	.855
45	.272	.475	.573	.544	.672	.810
60	.239	.448	.532	.478	.632	.755
75	.200	.405	.495	.400	.572	.700
90	.159	.354	.450	.318	.500	.635
105	.118	.290	.407	.236	.410	.575
120	.080	.222	.360	.160	.313	.512
135	.0466	.152	.307	.093	.214	.440
150	.0213	.085	.250	.043	.121	.355
165	.0054	.033	.165	.011	.0434	.250
180	0	0	....	0	0	.

full-wave rectification on single-phase alternating current. The data of this table are partially summarized in convenient form in the curves of Fig. 23. These curves show, for instance, that a firing angle of 60 degrees is required to reduce the average current by 25 per cent and 120 degrees for a 75 per cent decrease. If suitable instruments are located in the load circuit to determine the average and effective current, the firing angle can be determined approximately. Thus a 0.5 ratio of average to effective current will indicate a 120-degree firing

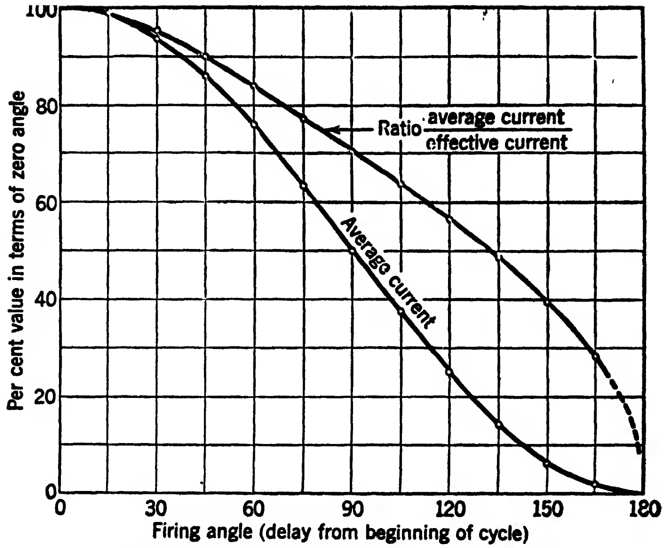


FIG. 23. Effect of variation of firing angle in thyratron circuits.

angle. The use of an oscillograph will give a more direct and accurate answer.

Mercury-Pool Tubes. A mercury-pool tube uses a pool of cold mercury for a cathode. The conducting medium consists of emitted electrons plus ions produced in the mercury vapor. The conduction due to the mercury vapor is of the arc-discharge type. Since electrons are not readily extracted from a cold metal, some special means must be employed (1) to initiate and (2) to maintain an emission of electrons from the mercury pool. Two general methods are employed for initiating the electron emission, one mechanical and one electrical. In the application of mechanical methods an auxiliary anode is caused to touch the mercury pool for an instant. This may be accomplished (1) by tipping the tube so that the liquid mercury flows to a point where cathode and anode are brought in contact, (2) by moving the anode so as to dip into the mercury pool, or (3) by squirting the liquid mercury upward so that it contacts an auxiliary anode momentarily. In any of these processes the momentary contact between the mercury (cathode) and anode causes a transient current followed by a spark and perhaps a small arc which results in an initial emission of electrons from the mercury. One electrical method for starting electron emission is to place an auxiliary anode (called an ignitor) so that its point dips into the mercury. Then a transient current from point to mercury

produces a spark for starting emission. In some mercury-arc lamps a very high transient electric field has been created close to the mercury pool by an inductive voltage "kick." This transient field has been sufficient to cause a breakdown and start of emission on the glow-discharge principle. After the emission of electrons is initiated the maintenance of emission depends on additional theoretical considerations. As soon as electron emission is initiated, the electrons are attracted toward the main anode and produce ionization. An arc develops and the plasma of positive ions fills most of the space between the anode and the mercury pool. The current seems to originate at one or more hot spots on the mercury pool. The hot spots look like little craters on the surface of the pool and they travel about over various paths or eddies. The craters are formed by the bombardment of the positive ions on the mercury. At first it was assumed that the temperature of the hot spots was sufficient to produce thermionic emission from the mercury. This concept has been abandoned because the temperature necessary for thermal emission would vaporize all the mercury and that does not happen. Accordingly, the theory that has rather general acceptance is that the positive ions form a very high positive space charge just above the mercury. The resulting sheath between the mercury and plasma is so thin that the electric field in this sheath is sufficient to secure emission through the principle of high-field emission.

One important advantage of all mercury-pool tubes is that the mercury cathode is capable of furnishing enormous emission temporarily without damage. Thus the tube will withstand temporary overload and even short circuit without destruction of the cathode. The mercury vapor condenses on the walls of the tube and returns to the pool.

One type of mercury-pool rectifier has been used since the early part of this century. This device and its circuit are shown in Fig. 24. The tube was built with two anodes for full-wave rectification. The glass enclosure had a large upper chamber to provide cooling area. The arc was started by tilting the tube so that the mercury could contact the auxiliary starting anode. These tubes had a maximum capacity of about 30 amperes and were widely used for charging storage batteries. A few tubes of this type are still in service for rectifying alternating current for use on series d-c street-lighting systems of the luminous-arc type. These early pool-type rectifiers had four disadvantages: (1) They had a fragile glass envelope; (2) they were not suited to the dissipation of heat; (3) they were limited in current capacity; and (4) their rectified voltage and current waves had wide fluctuations and were not suitable where a smooth flow of d-c power was

desired. These disadvantages led to the development of the water-cooled metal-tank multielectrode rectifier.

Multielectrode Metal-Tank Rectifier. The multielectrode rectifier was introduced into the electric-power field about 1925. This rectifier consists of a large cylindrical steel tank surrounded by a water-cooled jacket (Fig. 25) and evacuated to a pressure of one-millionth of an atmosphere. The cathode consists of a pool of mercury at the bottom (Fig. 26). There are from six to eighteen main anodes, like the one shown at the left of the figure. It consists of a graphite cylinder connected to the outside through an insulating bushing in the wall of the tank. The main anodes are placed in a cylindrical insulating shield. The value of using multielectrodes is twofold. First, electricity is transmitted most economically by multiphase circuits (usually three-phase), and, secondly, the multiphase rectifier gives a smoother output. The latter statement is illustrated by the rectified wave forms of Fig. 26. The single-wave, three-phase circuit gives a more desirable output wave, and the additions of more phases and anodes will increase the number of ripples but reduce their magnitude. A three-phase supply may be split into six, twelve, and eighteen phases by suitable transformer connections. The steel-tank, mercury-arc rectifier is a very rugged device, effectively cooled by water for any desired capacity and designed to give a smooth voltage, current, and power output.

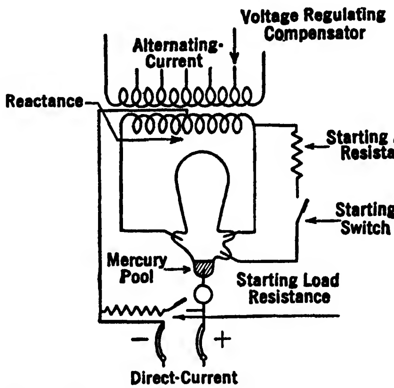


FIG. 24. Full-wave mercury-arc rectifier.

The steel-tank rectifier is started by the small central starting anode (Fig. 26) which can be lowered into the mercury and then withdrawn by an electromagnet. The breaking of the mercury contact on withdrawal strikes the arc which is then "picked up" by one of the main or auxiliary anodes which has the highest positive potential at that instant. The various main anodes will pick up the arc in rotation as they become sufficiently positive and maintain it until the load is taken over by another of higher positive potential. The output current at any instant is the arithmetic sum of the current of all anodes. Theoretically, only one anode having the highest positive potential would

be expected to carry the load current at a given instant. In practice the inductance in transformer windings in series with the anodes causes an anode to carry current for a time after a succeeding anode begins to conduct. Hence in practice, two or more of the multianodes

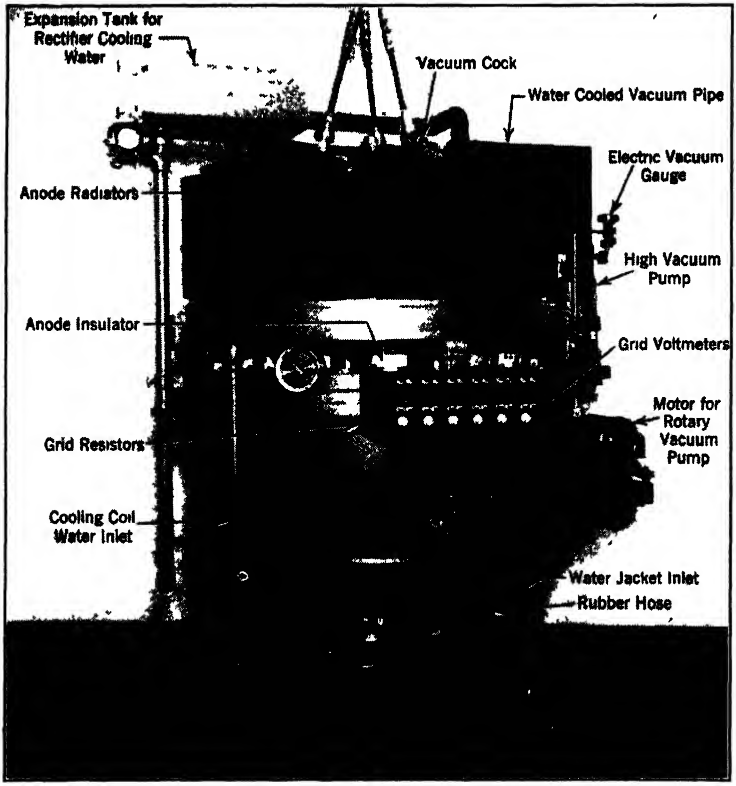


FIG. 25. Steel-tank mercury-arc rectifier. (Courtesy Allis-Chalmers Manufacturing Company.)

may be carrying part of the load current simultaneously. In the use of the multielectrode rectifier, the rectified d-c load may go to zero, which would extinguish the arc and make it necessary to re-establish ionization whenever the load returns. To avoid this contingency, one or more auxiliary anodes (see right side of Fig. 26) are constantly energized through a circuit independent of the load and they serve to keep gas ions present in the tank continuously. The continuous presence of gas ions in the arc chamber introduces a problem in the operation of the multielectrode rectifier. When any one of the multianodes

has a negative inverse voltage on it, it repels electrons and, theoretically, its current goes to zero, but positive ions may be present near it due to electrons en route to an adjacent anode. These positive ions will be attracted to the first anode and their bombardment may produce a hot spot on the anode which will emit electrons. Then this anode becomes a cathode and a reverse arc starts which will become a

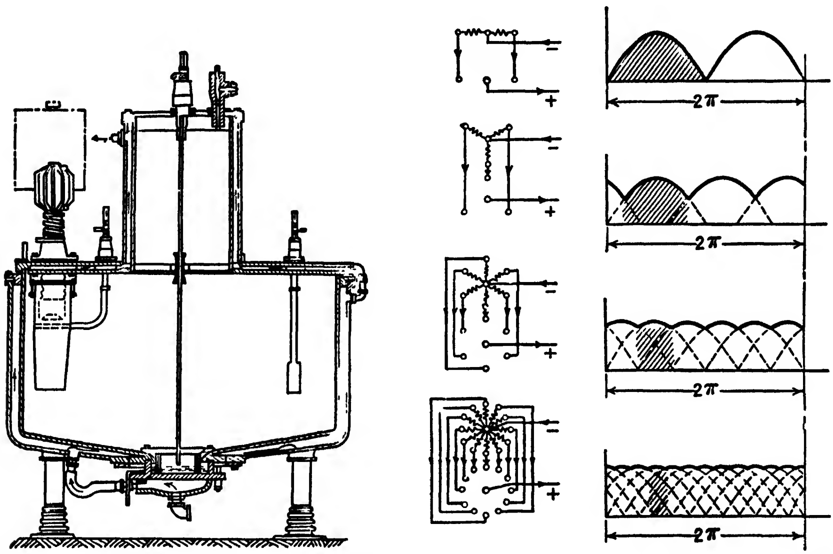


FIG. 26. Cross section of a steel-tank mercury-arc rectifier and rectified wave forms from simple and multiphase rectifiers. (Courtesy Allis-Chalmers Manufacturing Company.)

short circuit from anode to anode in the rectifier. Such a short circuit is called an "arc-back" or a "flash-back" and it will open the protective equipment and cause a shutdown.

The tendency toward arc-back is greatly reduced by the presence of a cylindrical shield placed around each anode. When the anode is positive, positive ions may be formed within the shield. However, as the potential falls to zero the positive ions are drawn out of the shield toward the cathode pool at the bottom of the tank or deionized by the shield. Hence the space within the anode shield becomes deionized and there is little tendency for arc-back. Occasional arc-backs do occur in the multielectrode mercury-arc rectifier.

The principle of grid control as used in other electronic devices may likewise be employed in the multianode rectifier. Here the grid is

placed within the grid shield chamber and close to the anode (see Fig. 26). In this position the grid can control the time of the starting of electron flow to the anode on each positive voltage wave. A negative voltage on the grid prevents starting, and the time when the grid potential changes to zero or positive may be controlled through a phase shift of the voltages supplied to the rectifier. This grid action controls the rectified output voltage and may be used to control output current for special applications.

The multielectrode steel-tank rectifier is widely used for large-power d-c applications. Over 3,000,000 kilovolt-amperes of these rectifiers are in service. The important applications are electric railway service, electrolytic processes in industry, and motive power in steel mills. The advantages of these units over rotary types of conversion are (1) simple and rugged construction, (2) long life, (3) high overall conversion efficiency, (4) high momentary overload capacity, and (5) quiet operation. Certain disadvantages of the multielectrode tank rectifier are causing it to be superseded by the single-anode type described on the following pages. These disadvantages are: (1) There is a greater possibility of arc-backs due to several anodes in the same arc chamber. (2) The multianodes require large spacing between cathodes and anodes with correspondingly higher arc drops and reduced efficiency. (3) The damage to any anode or part causes the entire unit to be taken out of service for repairs.

Excitron Single-Anode Rectifier. The excitron is a single-anode rectifier embodying the same general principles of construction and operation as outlined for the multianode unit. A cross section of an excitron unit is shown at the top of Fig. 27. The unit is placed in operation by a small jet of mercury thrown upward upon the excitation anode by the arc-starting device. As the mercury falls away a pilot arc is initiated and maintained by the d-c supply from the selenium rectifier and control circuit illustrated in the lower part of Fig. 27. The anode is surrounded by a grid consisting of a perforated graphite cylinder which serves to time the period of current rectification. Grid action also combines with the restraining influence of baffles to prevent arc-backs due to inverse voltages. Arc-back from one main anode to another (as in multianode rectifiers) cannot occur.

Excitrons may be used in groups of three, six, twelve, or eighteen as required to serve the same as the multianode units. The individual units are rather light in weight and are easily installed and repaired. Since the anodes are closer to the cathodes the arc drop is

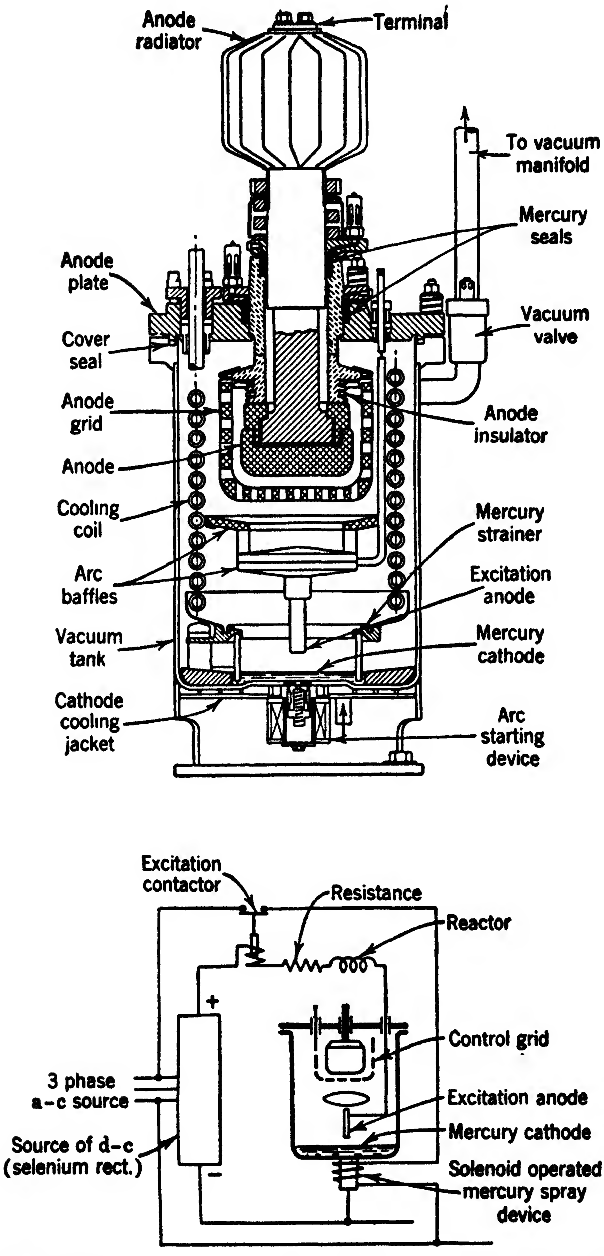


FIG. 27. Cross section and circuit of an excitron. (Courtesy Allis-Chalmers Manufacturing Company.)

lower and the efficiency is from 3 per cent to 4 per cent higher on output potentials of 250 to 300 volts. •

Ignitron. The ignitron is a type of three-electrode mercury-vapor tube developed by Slepian and Ludwig in 1932. The name ignitron was derived from the novel method of igniting or starting the arc in this mercury-vapor tube. The ignitron has a graphite anode in the form of a disk or cylinder. Its cathode consists of a pool of mercury placed at the bottom of the tube. The third electrode, called the *ignitor,* terminates in a point which dips into the pool of mercury (Fig. 28). The ignitor may be a rod of ceramic material such as silicon carbide or one having a graphite center with a boron carbide coating.

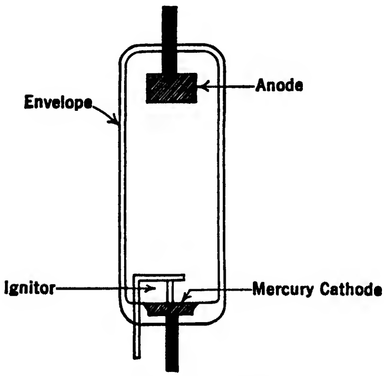


Fig. 28. An ignitron.

The theory of operation of the ignitron is simple. With a voltage placed across the cathode-anode circuit (anode positive) nothing happens because there are no electrons emitted from the cold mercury cathode. However, if a flash of current is passed through a circuit connecting the ignitor and cathode a spark will be created at their contact. This spark will produce some emission, and ionization of the mercury vapor will result. The arc thus established will continue as long as a suitable potential is maintained in the anode-cathode path.

The phenomena accompanying the development of the arc in the ignitron consist of two parts. The mercury does not come into intimate contact with the rough surface of the ignitor which consists (microscopically) of a number of sharp points. The flash of voltage across the ignitor-cathode circuit causes a potential gradient of approximately one million volts per centimeter across the point contacts, which is high enough to pull electrons out of the cold mercury. Simultaneous with this initiation of electron emission, the rise of current through the high resistance rod of the ignitor produces a potential gradient between the surface of the mercury and the top of the rod. This electric field accelerates the emitted electrons upward to ionization, and the development of an arc to the ignitor terminal which, in turn, is transferred upward to the anode (if positive).

The ignitron finds its application as a controlled rectifier for alter-

nating currents. Since the rectified anode current will go to zero each time the negative voltage loop is applied, it is necessary to ignite the tube for each positive cycle. This result can be obtained by the circuit given in Fig. 29. The hot-cathode diode in the auxiliary circuit conducts current as soon as its anode reaches the ionizing potential. This current passes through the ignitor, causing it to fire the ignitron since its anode becomes positive at the same time. When the ignitron fires, its arc-drop or cathode-anode potential falls to a low value of about 12 volts. This lowers the voltage across the series auxiliary

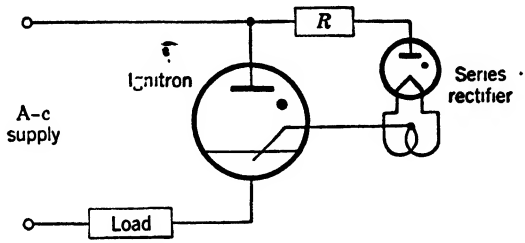


Fig. 29. Anode firing circuit for an ignitron.

circuit, consisting of rectifier, series *R*, and resistance at ignitor, so that the current in this circuit becomes very low or falls to zero.

The output of the ignitron may be regulated by controlling the time phase of firing or igniting. One method for accomplishing this control substitutes a thyratron for the diode rectifier in Fig. 29. The desired time of firing the ignitron is secured by either amplitude control or phase-shift control of the thyratron. Another method of firing and controlling the output of the ignitron utilizes special electrical networks having "trigger" action.

Ignitrons are often fired by the discharge of a condenser using some form of trigger circuit for initiating the discharge. One reactor-excitation circuit of this type is shown in Fig. 30. This circuit consists of three main parts: The firing circuit proper, a voltage-compensating network, and a phase-shifting reactor. In the firing circuit a capacitor *C* is charged through a linear reactor (constant reactance throughout range of applied voltage). The capacitor voltage is impressed across the ignitor circuit through a saturable reactor and directional filters (copper oxide rectifiers, page 318). The saturable reactor becomes magnetically saturated when the impressed voltage (and resulting current) reaches a predetermined value. At this critical point the reactance of the saturable reactor decreases and the capacitor discharges

a peak current through the ignitor-cathode circuit along the path shown by the arrows.

The function of the voltage-compensating network (C_1 and L_1 in parallel) is to furnish a nearly constant input voltage to the firing circuit regardless of fluctuation in the a-c supply. L_1 is a saturating reactor so designed that at normal line voltage its lagging circuit is just balanced by the leading current of capacitor C_1. If the line voltage is too high, the reactor L_1 begins to saturate and draws an excess mag-

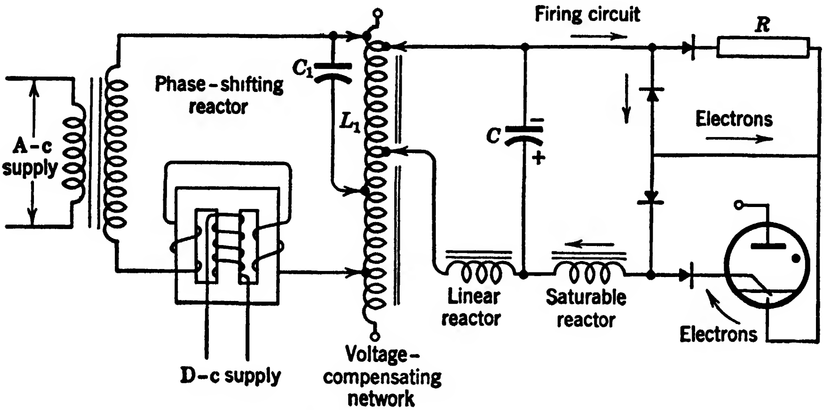


Fig. 30. Network for firing ignitrons (bold-face arrows show direction of conventional current flow).

netizing current. This lagging current passing through the phase-shifting reactor (left) produces a quadrature voltage drop and reduces the firing-circuit voltage. If the line voltage is low, L_1 draws a small magnetizing current and the stronger leading capacitor-charging current in passing through the phase-shifting reactor increases the firing-circuit voltage. This voltage-compensating action is so effective that the supply-line voltage can vary 50 per cent without ignition failure.

The phase-shifting transformer (often called a saturable reactor) has a three-legged iron core with series additive windings on the outer legs. With the d-c coil unexcited the series coils are highly inductive and will shift the phase of the voltage applied to the firing circuit. If sufficient direct current is passed through the coil on the central leg to saturate the entire core, the reactance of series windings falls to a low value and the phase of the voltage applied to the firing circuit is advanced. Intermediate values of direct current will give a corre-

sponding degree of phase shift which will determine the time of firing the ignitron.

The a-c supply for this reactor excitation must be the same as that applied to the anode-cathode circuit of the ignitron. The circuit of Fig. 30 will control single-phase, half-wave rectification. Full-wave rectification can be attained by substituting the ignitor-cathode circuit of a second ignitron for resistor R.

GENERAL CHARACTERISTICS

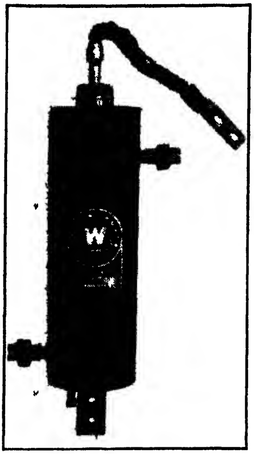
Mercury-pool triode	Clamp cooled
Line voltage range	250–600 volts
Tube voltage drop	12–18 volts
Ignitor voltage, max positive peak required	200 volts
Ignitor current, max peak required	30 amp
Ignition time, max	100 μsec
Net weight	1.5 lb
Shipping weight	3 lb
Maximum Anode Rating	
Demand at 12.1 average amp *	300 kva
Demand at 22.4 average amp *	100 kva
Time of averaging current at 250 volts	22 sec
Time of averaging current at 500 volts	11 sec
Current surge peak at 250 volts	3360 amp
Current surge peak at 500 volts	1680 amp
Maximum Ignitor Rating	
Voltage, max positive	900 volts
Voltage, max negative	5 volts
Current, max	100 amp
Current, 5-sec average	1 amp
Max temperature of cooling clamp	50° C

* Demand current and kva are on the basis of full-cycle conduction (no phase delay) whether or not phase control is used.

FIG. 31. Light-duty ignitron. (Courtesy Westinghouse Electric Corporation.)

The first ignitrons were encased in glass tubes as shown in Fig. 28. Later practice has tended toward building ignitrons in metal envelopes in order (1) to facilitate the dissipation of heat, (2) to increase their physical sturdiness, and (3) to improve the psychological reaction of industry toward the use of these tubes. The metal-encased ignitrons may be classified as to type of seal and type of cooling. In the smaller size the tube is evacuated and given a permanent seal (see Fig. 31). This type may rely on radiation and ordinary convectional circulation for cooling. The use of forced air circulation and also the addition of cooling fins will increase the rated capacity of these tubes. In the intermediate and large sizes the metal tubes are designed for cooling by water circulation through copper tubes or helical paths around the metal case (Fig. 32). In the large sizes the evacuation of the ignitron

is maintained by a mercury condensation pump. Although these larger tubes could be built with a permanent seal and thus avoid the cost and inconvenience of the evacuation equipment, another factor in economy demands the use of the nonseal type. This factor is the



IGNITRON WL-679

FOR RECTIFIER SERVICE

General Characteristics

Mercury-pool electronic tube—water cooled	
Tube voltage drop	12–20 volts
Ignitor voltage, max positive peak required	150 volts
Ignitor current, max peak required	40 amp
Ignition time, max	100 μsec
Temp rise, 150 amp average and water flow 1.5 gpm	7° C
Mounting position—vertical, cathode down	
Net weight	14 lb
Shipping weight	20 lb

Maximum Ratings 25 to 60 Cycles

Anode voltage, peak forward and inverse	900	2100 volts
Anode current, peak	900	600 amp
Anode current, continuous service, max average	100	75 amp
Anode current, 2-hour service, any 2-min average	150	112.5 amp
Anode current, 1-min service, any 1-min average	200	150 amp
Anode current, surge, 0.10-sec max	6000	4500 amp
Auxiliary anode voltage, peak inverse, main anode conducting	25	25 volts
Auxiliary anode voltage, peak inverse, main anode not conducting	150	150 volts
Auxiliary anode current, max average	5	5 amp
Ignitor voltage, max positive peak allowed	900	2100 volts
Ignitor voltage, max negative peak allowed	5	5 volts
Ignitor current, max peak allowed	100	100 amp
Ignitor current, max average 10-sec max averaging time	2	2 amp
Cooling water flow, min	1.5	1.5 gpm
Cooling water temperature max at outlet	60° C	45° C
Cooling water temperature min at inlet	10° C	10° C
Cooling water temperature range, optimum	20–45° C	20–40° C

For Welding Service

Line voltage, rms	2400 volts
Maximum demand at 75 average anode amp	1200 kva
Maximum demand at 113 average anode amp	600 kva

FIG. 32. Heavy-duty ignitron. (Courtesy Westinghouse Electric Corporation.)

uncertainty covering the life of the sealed unit and a preference for the opportunity to take the unit apart for repairs rather than to discard a very large and expensive device. An ignitron of large capacity is illustrated in Figs. 33 and 34.

It is logical to assume that the ignitron should withstand an infinite inverse voltage without arc-back. It would seem that all positive and negative ions would be swept out of the cathode-anode arc at the con-

clusion of the positive-voltage loop, and hence no inverse voltage could cause a breakdown. This logical theory was assumed by the early investigators of this device, but it has not been proved by practice. Occasional arc-backs do occur in ignitrons though they are not damaging to equipment. These arc-backs may be due to tiny hot spots which develop on the graphite anode and serve temporarily as emitters.

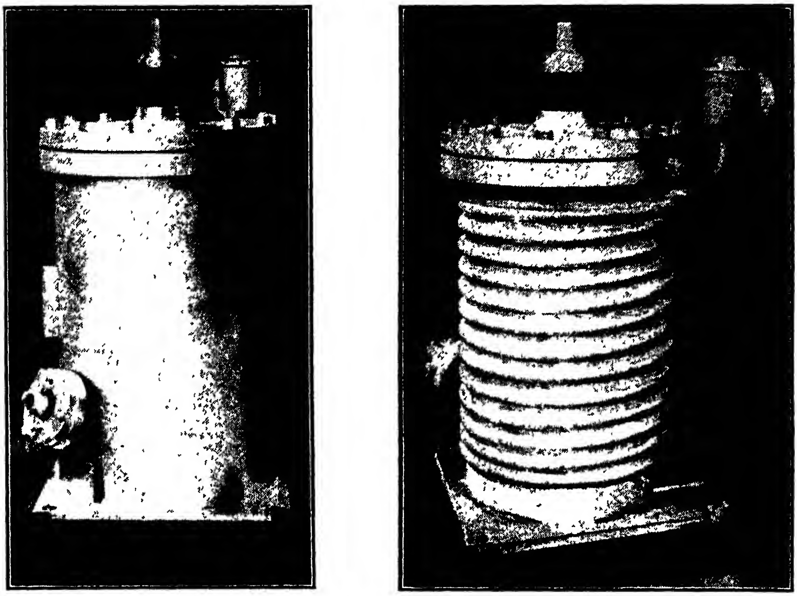


FIG. 33. Large-size water-cooled ignitron (with and without outer case). (Courtesy Westinghouse Electric Corporation.)

In order to improve the arc-back characteristic, baffles may be placed in the ignitron as illustrated in Fig. 34. By increasing the length of the arc path these baffles reduce the possibility of a reverse flow of current. An anode shield or shield grid placed around the anode with suitably timed potential will further serve to reduce any tendency to arc-back (Fig. 34).

There are two important applications of the ignitron in industry. The first is in the control of resistance welding—spot, butt, and line welding. The second application is for rectification for d-c power applications. The use of the ignitron has been rather revolutionary in making possible new applications of welding because it assures a uniformity of results never attained in the past. In this field a combination of two ignitrons serves as an electronic switch for controlling the

time of application of an alternating current in making resistance welds. A simplified circuit for performing this function is given in Fig. 35. Two ignitrons are connected with reversed polarity in parallel with each other and in series with the primary of a welding transformer. The two ignitors are connected in series with each other and with a

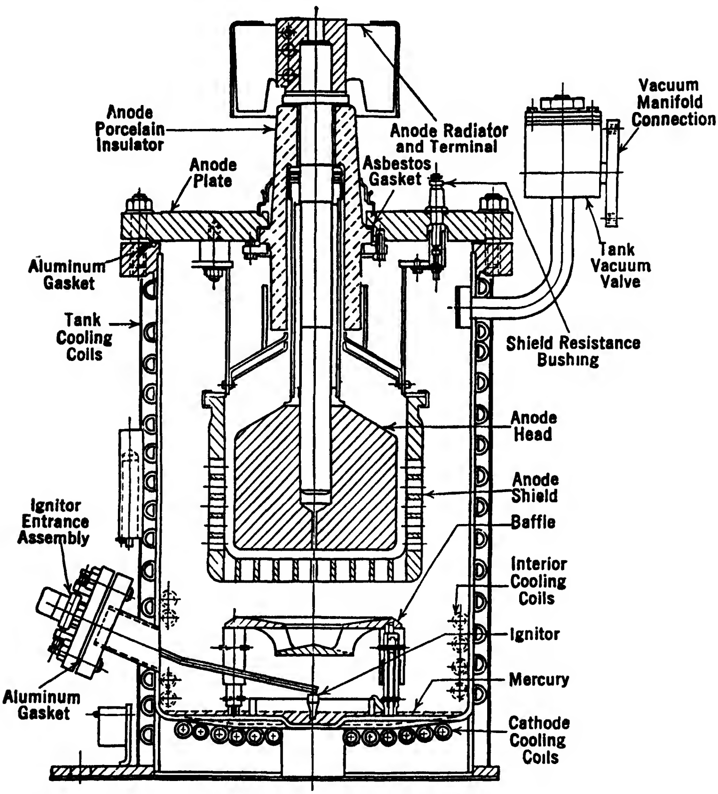


FIG. 34. Cross section of a large-capacity ignitron. (Courtesy Westinghouse Electric Corporation.)

timer contact. As long as the timer contact is "open," neither ignitor can be fired and the primary circuit is "open," being blocked by the cathode-anode gaps in the two ignitrons. The welding operation is started by the closing of the timer contact which permits current to pass through the two ignitors in series along the path indicated by the arrows. This current produces a spark at both ignitors and causes that tube having a positive anode to fire. After firing, the potential across this ignitron and the two ignitors in turn falls to a low value,

causing the current through the ignitors to drop to zero. As the a-c voltage rises on the next half-cycle, a flash of current will again pass through both ignitors and the other ignitron will fire. Thus the two ignitrons serve as an "electronic switch" to pass the alternating current through the transformer primary for operating the welder. Whenever the timer contact is opened the ignitrons again block the circuit.

Ignitron versus Thyratron. Although ignitrons and thyratrons have many characteristics in common, they are not competitive in their applications. The thyratron is essentially a low-current and high-voltage

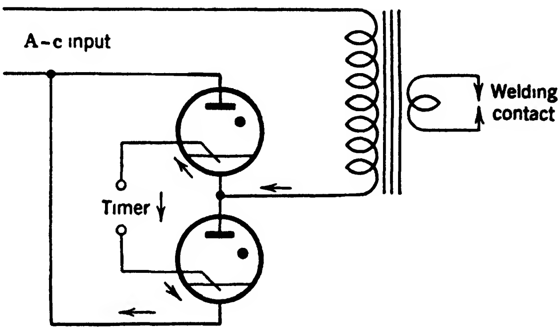


FIG. 35. Circuit for an electronic control switch consisting of two ignitrons.

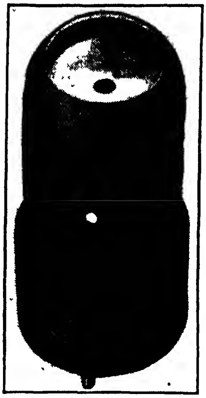
device. It can be made in small sizes for many small power applications. It has the advantage of greater simplicity of control circuits and the disadvantages of (1) the continuous use of energy for heating the cathode when it is not in use, and (2) the likelihood of damage from overload and overvoltage.

The ignitron is essentially a large-current and a low- or medium-voltage device. It is suitable for those applications having high peak current and high power requirements where the cost of the complicated ignition system does not form an excessive percentage of the total cost. The ignitron has the advantage of exceptional sturdiness, and it will withstand temporary heavy overloads and even short circuits which would ruin a thyratron. The ignitron does not require any energy for operating the cathode except that used for the ignitor when the tube is supplying rectified current. The energy for operating the ignitor is very small since the actual firing current flows for a few microseconds.

Capacitron. Another method of "firing" a mercury-pool rectifier uses an insulated conductor placed above the mercury pool. Either a

high-frequency field or a high-voltage surge impressed on this conductor will serve to control the time and firing of a pool tube. A device using this principle is called the *capacitron.* The device has little application at present.

Cold-Cathode Tubes. A number of tubes using a cold metal (other than mercury) for cathodes have been developed and employed for useful purposes. These tubes are filled with inert gas under low pressure and operate on the principle of the glow discharge. They serve (1)



Classification—recording lamp (cold cathode)	
Overall length, max	3¼ in.
Max diameter	1.275 in.
Mounting position—any	
Starting voltage, max	170 d-c volts
Operating voltage, max	135 d-c volts
Operating current	5–35 ma
Frequency range	15–15,000 cps
Useful light range	3500–6500 A

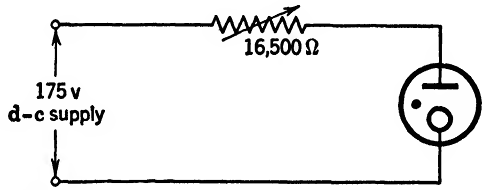


FIG. 36. Cold-cathode variable-light tube and circuit. (Courtesy Sylvania Electric Products, Inc.)

as sources of light, (2) for voltage regulation, (3) for rectification, and (4) as "trigger" or control tubes. The *glow tube* has two cold electrodes and gives forth light (negative glow) when the breakdown voltage is applied (see Fig. 8, page 258). If a d-c voltage is applied the glow appears on the surface of the cathode, but if a-c voltage is impressed the glow appears at both electrodes since each serves alternately as a cathode. With simple flat electrodes lying in the same plane and with a constant source of potential the tube serves as a constant source of light and may be used as a signal light, pilot light, and test light. The source of light between two electrodes may be varied by changing the geometry of the electrode construction. Thus if the cathode is a small solid cylinder closely surrounded by a hollow cylinder, the light appears at the ends of the electrodes. The intensity of the light produced will vary as the value of the current conduction. Thus with proper construction a glow tube may become a modulated light source. Such tubes have been used for recording sound on film (light on sound track), for early types of television receivers, and for

facsimile transmission today. A commercial tube of this type, its rating, and its circuit are illustrated in Fig. 36.

For high-speed photography and stroboscopic work special types of glow tubes have been developed. One of these tubes (Fig. 37) utilizes

Classification—Strobotron cold cathode	
Physical	
Overall length	$4\frac{9}{32}$ in
Seated height, max	$3\frac{5}{8}$ in.
Diameter, max	$1\frac{3}{16}$ in.
Mounting position—any	
Electrical Design Characteristics	
Anode voltage	350 volts, d c
Average anode current, max	50 ma
Instantaneous anode current, min	5 amps
Grid No 1 d-c voltage, max	70 volts
Grid No 2 d-c voltage, max	70 volts
Grid current (max average either grid)	15 ma
Electrical Operating Characteristics	
Grid No 2—cathode starting voltage	80–125 volts
Pulse frequency range	240 pps
Tube voltage drop	
Glow discharge	75 volts
Arc discharge	20 volts

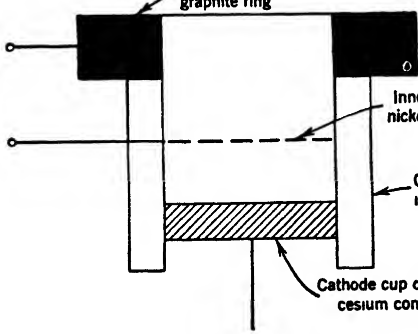
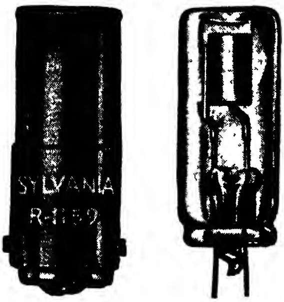


Fig. 37. Stroboscopic cold-cathode light source. (Courtesy Sylvania Electric Products, Inc.)

a double-grid structure which permits quick firing and gives brilliant flashes lasting for only a few microseconds. This tube was developed by Germeshausen and Edgerton. The cathode cup contains a cesium compound that liberates cesium at a relatively low temperature. Breakdown is initiated between two of the electrodes, grid to cathode, or grid to grid. Application of this tube is covered in Chapter XVI.

One important characteristic of the glow-discharge tube is the nearly constant voltage drop for a wide range in current variation. This property is utilized on a form of glow tube known as a *voltage-regula-*

tor tube. A commercial regulator tube of this type is illustrated in Fig. 38. Larger tubes having constant operating voltages up to 150 volts



Length	1¾ in.
Diameter	⅝ in.
Bulb—metal enclosure	
Starting voltage, max	170 d-c volts
Regulating voltage	65–75 volts
Operating current	0.5–5.0 ma
Regulation	±1.5 volts

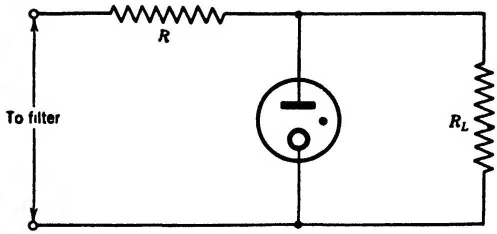


FIG. 38. Miniature voltage-regulator tube, rating, and circuit. (Courtesy Sylvania Electric Products, Inc.)

and currents up to 40 milliamperes are available. An early type of cold-cathode rectifier was known as the Raytheon BH (Fig. 39). It was a full-wave rectifier having a metal shell (an inverted metal bowl) for a cathode and two anodes consisting of insulated metal wires, the ends of which projected into the bowl. The tube was designed to utilize Paschen's law (see page 254). Thus the cathode-to-anode spacing gave the minimum voltage for ionization, whereas the short anode-to-anode spacing permitted a very high voltage for the gas pressure used. Thus conduction does not take place from anode to anode. The normal cathode-to-anode voltage drop was rather high and the operation was very sensitive to the gas pressure. Changes in gas pressure due to leaks or clean-up made the tube inoperative. Hence this device was withdrawn from the market in the early thirties and has been replaced by hot-cathode rectifiers.

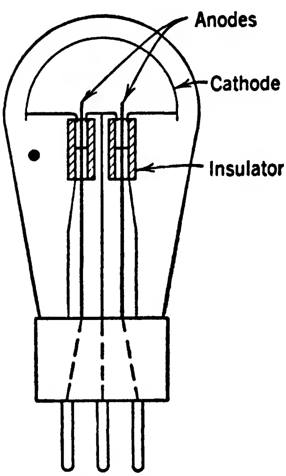


FIG. 39. Obsolete cold-cathode full-wave rectifier tube.

Another type of cold-cathode, grid-glow rectifier tube finds wide application in subscriber's telephone apparatus where its use simpli-

fies selective ringing on four-party lines. The essential parts of this tube are shown in Fig. 40. The tube employs two semicircular bell-shaped disks as cathodes. Between these cathodes is placed an anode consisting of a small circular rod enclosed in a glass tube. The tube contains neon gas under low pressure. A nominal breakdown potential of 70 volts is required to create a glow discharge from cathode to cathode, but because of greater spacing a breakdown potential of 175

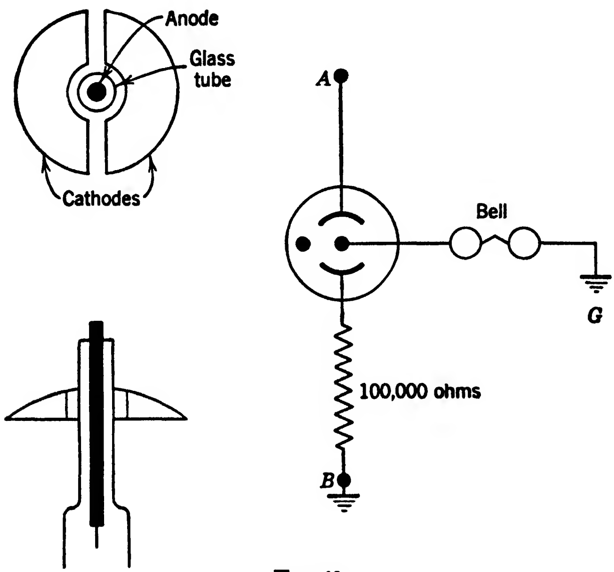


FIG. 40

volts is necessary for a like discharge from either cathode to the anode. However, while a glow discharge exists between cathodes, a potential of only 75 volts is necessary for unidirectional electron conduction from either cathode to the anode.

This tube is used in the circuit shown on the right of Fig. 40. To ring the bell a pulsating unidirectional potential is impressed between the line A and ground. This voltage causes a glow discharge between cathodes with a minute current through the high resistance and simultaneously a pulsating unidirectional current through the telephone bell. The pulsating unidirectional voltage is produced by superimposing an a-c voltage upon a battery supply. Individual party ringing is effected because the bells are selective with respect to the direction of current pulses and are properly distributed between line wires and ground. One such tube is shown in Fig. 41.

One type of gaseous rectifier starts with a cold cathode and operates on thermionic emission. The tube contains inert gas and has an oxide-coated cathode without a heater. With a minimum peak anode supply potential of 300 volts, a glow discharge takes place and the resulting bombardment of the oxide-coated cathode raises its temperature to the point necessary for satisfactory thermionic emission. After the cathode is heated the cathode-anode average potential drops to about 24 volts. This voltage is higher than that of a hot-cathode type with heater because the energy for heating the cathode is developed in the cathode-anode circuit. These tubes are known as *ionic-heated cathode* rectifiers. The output current must be limited by a suitable load impedance and a minimum cathode-anode current is required to keep the cathode hot. Examples of the ionic-heated cathodes are the OZ4 and the OZ4-G. The principle of the ionic-heated cathode is employed in the operation of some types of fluorescent lamps.

FIG. 41. Cold-cathode relay tube. (Courtesy Western Electric Company.)

The *grid-glow tube* * is a gas-filled, three-electrode tube having a cold cathode. The tube functions somewhat like the thyratron through the trigger action of its grid. Since it has a cold cathode the tube conduction is of the glow-discharge type involving much higher anode and grid voltages. These higher voltages mean a higher percentage of loss in the tube and lower rectification efficiency. However, the cold cathode does not consume any energy even though the tube is connected in a circuit continuously. The grid-glow tube is constructed much differently from a thyratron, as shown in Fig. 42. The anode is a wire encased in a glass tube and placed at the center of a much larger cylindrical tube which constitutes the cathode. A metal tubular shield surrounds the glass tube and anode. The grid is a small ring surrounding the projecting top of the anode. The anode is usually operated at a voltage above the breakdown voltage for the glow discharge of the gas from anode to cathode but below the breakdown for grid-to-anode spacing (see page 253) If the grid is negative or of low potential, most of the lines of the electrostatic field terminate on the grid and no breakdown from cathode to anode can occur. The usual

* The term grid-glow tube has been applied to the thyratron by some writers in the past. The usage here is in accord with current standards.

type of circuit for utilizing the grid-glow tube is shown in Fig. 43. The shield is connected to the cathode outside the tube by a resistance of the order of 5 to 10 megohms. The control grid potential is adjusted by the values of R and C. The load in the anode circuit is usually a relay. Changes in R or C or even in the value of the impressed a-c or d-c voltage will serve to fire the tube. Change in the capacity of C may be produced by the capacity of an approaching human hand or some other object. A photocell may be substituted for R or C whereby a change in light will fire the tube.

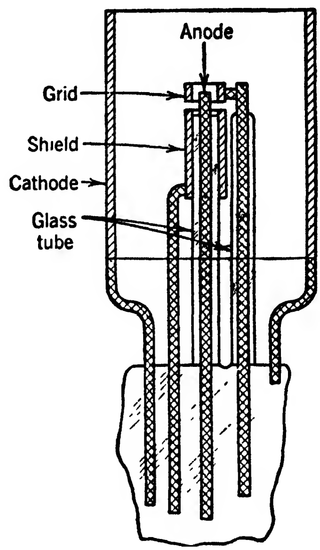


FIG. 42. Construction of a grid-glow tube.

The cathode of the grid-glow tube is indestructible but too high a voltage on the anode will cause the glow discharge to change to an arc and thus burn out the tube.

The grid-glow tube is used where a very sensitive control is desired to operate a relay and where the current delivered by the tube is sufficient to meet the requirements of an application. A commercial grid-glow tube, its rating, and its characteristics are given in Fig. 43.

Effect of Gas in a Triode. The discussion of the high-vacuum triode in Chapter IV ignored the presence of a small amount of gas which remains after the best evacuation processes are used. An interesting phenomenon produced by the presence of this gas may be observed if delicate measurements are made of the grid current under varying grid potentials and plotted to an enlarged scale. This phenomenon is illustrated in the curve of Fig. 44, which shows a reversal of the direction of grid current in the region of zero grid voltage. Some of the gas molecules in this high-vacuum triode are ionized by bombarding electrons and the positive ions thus formed are attracted to electrodes having a negative potential. When the grid is made positive, part of the electrons emitted by the cathode will be attracted to the grid and will constitute the positive grid current as shown by the curve and by the full-line arrow in the circuit diagram of Fig. 44. Now as the potential on the grid goes to zero, or negative, the grid will attract positive ions and the positive ions landing on the grid will take electrons

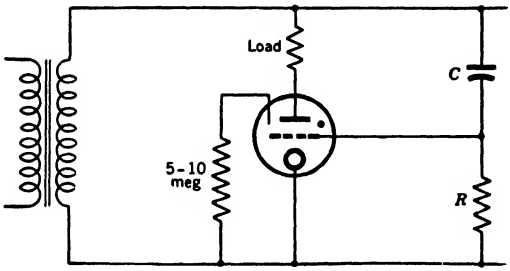
GRID-GLOW TUBE

General Characteristics

Air-cooled tetrode	
Cold cathode	
Ionisation time	10 μsec
Deionization time	1000 μsec
Tube voltage drop, average	180 volts
Control characteristics—positive and negative	
Mounting position—any	

Maximum Ratings

Anode voltage, peak forward	500 volts
Anode voltage, peak inverse	800 volts
Anode current, average	0.015 amp
Anode current, peak	0.10 amp
Anode current, surge, for design only	1.0 amp
Grid voltage, peak, positive	300 volts
Grid current, average	0.2 ma
Grid current, peak	0.8 ma
Averaging time, anode and grid currents	10 sec
Temperature range	−40° to +70° C



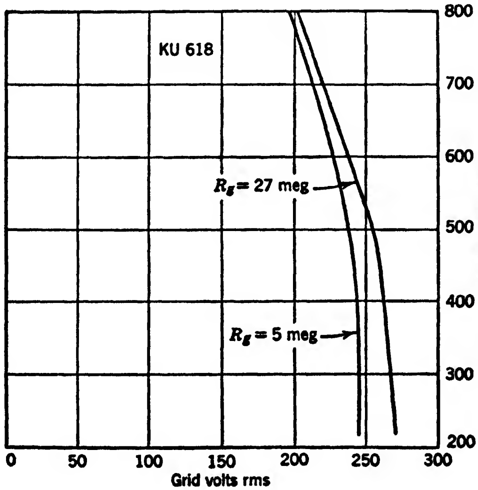


FIG. 43. Grid-glow tube characteristics, rating, and circuit. (Courtesy Westinghouse Electric Corporation.)

from the grid and become neutral atoms. Since the grid now has no internal source of electrons to neutralize the positive ions, electrons must come through an external circuit leading from the cathode or anode through the grid circuit to supply this need. Hence the current in the grid circuit reverses or becomes negative as shown by the curve and by the dotted arrow. As the grid is made strongly negative the electrons emitted are blocked from passing to the anode and ionization ceases. This cessation of ionization would be expected to reduce the grid current to zero, yet at this juncture the magnitude of the negative grid current begins to increase along the linear path indicated by the dotted line.

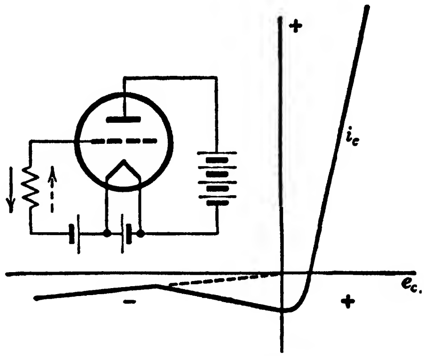


FIG. 44. Grid current in a high-vacuum triode.

This reversed current is a minute leakage current produced by the negative grid potential impressed across the high (usually tens of megohms) resistance between the glass-insulated leads to the grid and cathode. It should be noted that the reversed grid currents are very small in magnitude and are not indicated by ordinary meters. Hence this phenomenon is of theoretical interest primarily though it may assume importance where measurement of infinitesimal magnitudes are involved.

PROBLEMS

1. The mercury-vapor rectifier tube of Fig. 6 is connected in series with a supply of 115 volts a-c and a load resistor. If the arc drop is assumed to be 15 volts and the peak pulse current is 0.5 ampere, what is the minimum value of load resistance to protect the tube from overload on continuous service?

2. The phanotron of Fig. 7 is connected for half-wave rectification in series with a load resistor to a 230-volt a-c supply. The rated average load current is 20 amperes which is 0.318 times the peak of the current pulse. Assume an arc drop of 15 volts and calculate (a) the magnitude of a load resistor to permit full load continuously.

(b) Now assume that, with the calculated resistor in the circuit and with the cathode only partially heated and giving a saturation emission of only 2 amperes, the supply-line switch is closed. What current will flow? What voltage exists across the cathode-anode circuit? What will happen to the phanotron?

3. (a) The thyratron of Fig. 19 is connected in series with a load resistor and 115-volt a-c supply line. What should be the value of the load resistance to limit the current to the rated average value?

(b) Assume that while the thyratron is cold and there are zero volts on the grid the supply line is connected to the cathode-anode circuit followed by a closing of the circuit for heating the filamentary cathode. What will happen?

4. The thyratron of Fig. 19 is in normal operating condition and a d-c voltage of 400 volts is to be applied to its anode. What voltage must be applied to its grid to prevent firing? Give the range of grid voltage values necessary to cover the complete range of uncertainty arising from ambient temperatures.

5. Assume that the thyratron of Fig. 19 is to be fired by an a-c voltage which rises to 30 volts + maximum and that the arc drop from cathode to grid is 10 volts. What should be the value of a resistor placed in series with the grid to limit the grid current to 0.11 ampere?

6. The thyratron of Fig. 20 is to be used for half-wave rectification of an a-c voltage peaking at 6200 volts, 60 cycles. A grid bias of −10 volts is applied to the tube. What should be the magnitude of the load resistance to limit the peak current to the rating shown?

If it is desired to fire the tube with a lag of 90 degrees behind applied voltage by superposing an in-phase a-c voltage in series with the grid (−10-volt bias), what should be the peak value of this a-c grid voltage? Use the middle curve of the characteristic. Is this particular condition for firing desirable? Discuss.

7. Construct a circuit for using the shield-grid thyratron of Fig. 21 for half-wave rectification with an anode peak voltage of 2000. Indicate the magnitude of voltages and currents at all parts of the circuit. Take values from ratings given in figure.

8. Construct a simple circuit showing how an ignitron may be fired by a thyratron using the discharge from a condenser.

9. The strobotron of Fig. 37 is flashed by the discharge of a 4-microfarad condenser (charged to 200 volts) in 100 microseconds. How much power does this represent in watts?

10. The thyratron of Fig. 19 is to be operated in a circuit having an impressed a-c sine-wave voltage of 710 rms. What negative bias voltage should be applied to the grid to cause the tube to fire with a phase-angle lag of 30 degrees, 45 degrees, 60 degrees, 90 degrees?

11. In the phase-shift circuit of Fig. 15, R has a value of 400 ohms and L of 500 millihenries. Calculate the angle of a phase shift on a 60-cycle circuit.

12. In the phase-shift circuit of Fig. 15, R has a value of 4000 ohms and C of 2 microfarads. Calculate the angle of phase shift on a 60-cycle circuit. (See text for relationship of R and C in shift circuit.)

13. What should be the extreme values of R in Problem 12 to give phase shifts of 20 degrees and 75 degrees?

REFERENCES

HULL, A. W., "Hot-Cathode Thyratrons," *Gen. Elec. Rev.*, **32** (1929).

WINOGRAD, H., "Development of Excitron-Type Rectifier," *Trans. Am. Inst. Elec. Engrs.*, **63** (1944).

SLEPIAN, J., and L. R. LUDWIG, "A New Method for Initiating the Cathode of an Arc," *Trans. Am. Inst. Elec. Engrs.*, **52**, 693–700 (1933).

GERMESHAUSEN, K. J., and H. E. EDGERTON, "A Cold-Cathode Arc-Discharge Tube," *Elec. Eng.*, **55**, 790 (1936).

Chapter XIII

CRYSTAL AND METALLIC RECTIFIERS

The phenomenon of nonlinear voltage-current relations in some substances and at the junctions of dissimilar substances has been known for a long time. As early as 1834, Faraday discovered the highly negative temperature coefficient of resistivity exhibited by silver sulphide. Early in this century many materials were employed in "wireless" reception which exhibited highly polarized nonlinear characteristics. These materials consisted of a lump of some mineral such as galena, silicon, iron or copper pyrites, zincite, bornite, or silicon carbide. In the process of signal detection these materials made up one side of a circuit while a feeler wire called a "catwhisker" bearing upon the material was connected to the other terminal. This combination known as a crystal detector had a unilateral current characteristic in which electrons moved from the crystal to the metal point bearing upon them. The disadvantages of this early device were that the crystal had "sensitive spots" which were destroyed by overload and heat and the sensitivity could be lost by slight displacements of the catwhisker. This form of crystal detector or rectifier was widely used in the amateur period of wireless telegraphy and the early days of voice radio. The more stable operation of the vacuum-diode detector caused the crystal detector to be displaced and fall into disuse shortly after 1920. Experimentation on the construction and use of crystals as rectifiers was resumed during the 1930's, and two new crystal rectifiers, silicon and germanium, developed during this period were put into effective use during World War II.

The unilateral current-carrying characteristic of a junction of selenium and a metal was discovered in 1883, though commercial application of the combination was not made until 1930. A similar discovery of the unilateral characteristic of copper oxide on copper was made about 1920 and resulted in an early application of this combination.

The crystal and the semiconductor rectifiers suggested in the preceding statements are not electronic devices in the narrow sense of that term since they do not involve movements of electrons in a vacuum or

ions in gases under low pressure, but they do serve as important components in electronic combinations and circuits.

Silicon and Germanium Crystal Rectifiers. The basic construction of the modern crystal rectifier is illustrated in Fig. 1. A tungsten wire having an S or a hook shape has a sharp point which bears on a semiconductor or rectifying material. The tungsten wire is fastened to one terminal and the semiconductor or crystal to the other terminal. The terminals are separated by an insulator (usually ceramic) and the intervening space may be filled with a gas, a plastic, or other suitable material. Metal end pieces or caps give mechanical strength and complete the construction of the crystal rectifier.

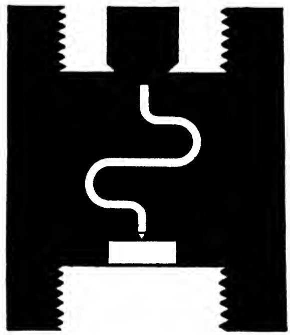


FIG. 1. Typical construction of a crystal rectifier.

When the tungsten terminal is made positive, electrons move in a forward direction from the semiconductor into the tungsten, but, when the polarity is reversed, very few electrons are able to cross the boundary into the semiconductor. The relationship of this unilateral conduction is shown in the curve of Fig. 4 which is typical for crystal and other semiconductor-to-metal rectifiers. This phenomenon of unilateral conductivity is not readily explainable.* The author offers the following. First, it should be noted that the contact between the metal and the conductor is a mere pressure contact, a simple type of boundary, and not a molecular bond. Second, it should be remembered that the contact is a point contact wherein the potential gradient surrounding the point will be high when differences of potential are applied. The electron energy levels which exist in the metal and semiconductor when separated are shown in part *a* of Fig. 2. Here the work function of the metal ϕ_t is relatively low compared to the total work function ϕ_{sc} for the semiconductor. Bringing the metal and semiconductor together in a mere physical contact without any applied potential will not produce much change in the energy levels. However, if a positive potential is applied to the metal, the high potential gradient close to its point will add to the normal electron energy level in the semiconductor near the contact and give a distribution as shown in part *b* of Fig. 2. Under this condition electrons can pass "over the hump" into the tungsten and this constitutes the forward current. When

* For one explanation see "Germanium Crystal Diodes," by E. C. Cornelius, *Electronics*, February 1946.

the polarity of the applied voltage is reversed, the negative potential (gradient) at the tungsten point lowers the energy level of the electrons in the semiconductor as shown in part *c* of Fig. 2, and very few electrons are able to pass to the semiconductor. The latter situation is somewhat analogous to an attempt to remove electrons from a cold electrode.

The crystal rectifier is small in size and light in weight. For use at ultra-high frequencies it is superior to the vacuum diode as a detector because (1) it has smaller input and output capacitances, (2) it does not introduce transit time effects, (3) noise voltages and interelectrode reactances have been reduced, and (4) it does not require any power for heating a cathode.

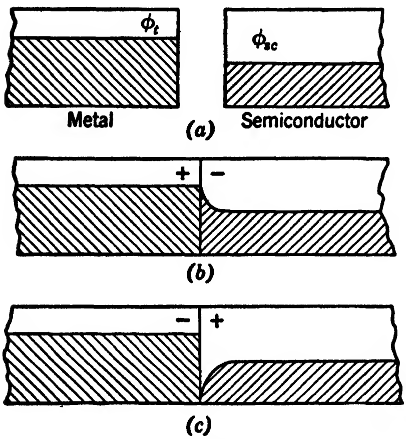


Fig. 2

One type of *silicon* crystal rectifier is illustrated in Fig. 3. It is approximately $1\frac{3}{16}$ inch long and has a maximum diameter of $\frac{5}{16}$ inch. This rectifier has been designed for easy insertion and plugging into cavity resonators and coaxial butterfly circuits. It is usable as a signal pick-up probe in wave guides and cavities. Silicon rectifiers are designed for use in the first detector stages of super-high-frequency superheterodyne receivers, and for tuned radio-frequency and video detector use at frequencies within the range of 1000 to 25,000 megacycles.

The *germanium* crystal rectifier, or crystal diode, as it is sometimes called, has a cartridge-type construction and the appearance of a one-watt fixed resistor as shown in Fig. 4. A tungsten feeler or catwhisker bears upon a thin disk of optically polished germanium soldered to the end of a wire. The germanium disk is prepared from the natural dioxide form (GeO_2) by reduction in hydrogen, leaving the amorphous metal in a pure state. The dull gray powder is melted and a small amount of tin is added. When it has cooled, a crystallized ingot results which is cut into wafers 0.6 millimeter thick and 3 millimeters square. The polished disks form a lattice-imperfection semiconductor.

The rating and typical characteristics of the selenium crystal diode

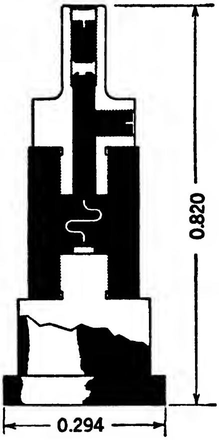
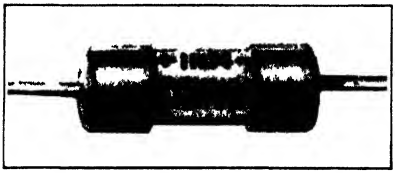


FIG. 3. Typical silicon crystal rectifier (Courtesy Bell Telephone Laboratories.)



RATING

Peak inverse anode voltage, max	50 volts
Peak anode current (sine wave), max	60 ma
Average anode current, max	22 5 ma
Surge current (transient peak), max	100 ma
Back conduction at 50 volts, max	2 ma

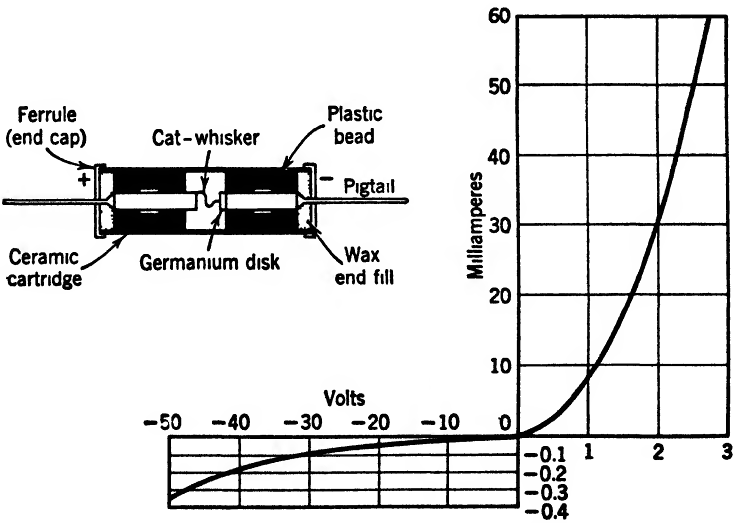


FIG. 4. Construction and characteristics of a typical germanium crystal diode. (Courtesy Sylvania Electric Products, Inc.)

are given in Fig. 4. This crystal diode has a shunt capacitance of approximately 3 micromicrofarads compared to approximately 15 micromicrofarads for a 6H6 vacuum diode under analogous conditions. This germanium crystal unit is recommended for use between 0 and 100 megacycles though it has been used up to 500 megacycles. It was designed for superheterodyne second detector applications, but it is useful in frequency modulation and in television reception where load resistances are low.

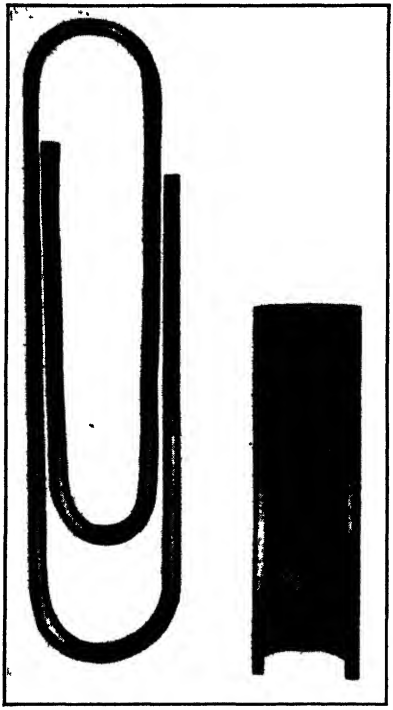


Fig. 5. Microphotograph of Transistor and paper clip. (Courtesy Bell Telephone Laboratories.)

The Transistor. A new device, the Transistor, which constitutes an extension of the principle of the germanium crystal diode was announced by the Bell Telephone Laboratories on July 1, 1948.* This device, *a crystal triode,* takes its name from the words TRANSfer resISTOR. Two point contacts (catwhiskers) consisting of 2-mil tungsten wires bear upon the upper surface of a germanium crystal and are separated by approximately 0.002 inch. This combination is encased in a metal cylinder $\frac{3}{16}$ inch in diameter and $\frac{5}{8}$ inch long, as shown in Fig. 5. The comparison of the size of the Transistor to a paper clip shows that the former is smaller than a subminiature vacuum tube.

Phosphor bronze wires have been used as well as tungsten, and any wire employed must have spring action to maintain a suitable pressure on the crystal. The spacing between the two catwhiskers is not too critical but is of the order of a few one-thousandths of an inch.

The Transistor may be employed as an amplifier or as an oscillator. The circuit for the Transistor used as an amplifier is given in Fig. 6. The action of either point electrode or catwhisker when used

* A more complete discussion of the Transistor may be found in the *Physical Review,* Vol. 74, pp. 230–233, July 1948, and in the September 1948 issue of *Electronics.* These articles report the findings of Dr. William Schockley, Dr John Bardeen, and Dr. W. H. Brattain, of the Bell Telephone Laboratories.

alone is that of a rectifier with the electrons moving from the semiconductor germanium to the point as explained in the preceding pages. When both electrodes are used simultaneously in the circuit of Fig. 6, a new and important action is obtained. The electrode on the left is called the *emitter* and corresponds to the grid of the vacuum triode. The emitter is given a positive bias (instead of negative) and it conducts electrons (rectifying action) in the direction indicated by the arrow. Obviously, the magnitude of this electron conduction is controlled by the applied voltage, and hence the signal voltage impressed in series in this circuit varies directly the electron current in the emitter circuit. The second electrode, called the *collector*, corresponds to the plate of the triode and has a negative (reverse) battery potential of approximately 50 volts applied to it. If the collector were used alone, this collector circuit would conduct very few electrons because of the blocking action at the catwhisker germanium crystal contact. However, with the combined emitter and collector circuits a different action takes place. Under the conditions of the combined circuits the emitter may be viewed as withdrawing electrons from the surface of the germanium in the region surrounding its contact with the surface. The fact that the withdrawn electrons leave "holes" in the semiconductor permits the excess of electrons at the near-by collector point to flow in and fill these holes. In this way the changes in signal voltage in the emitter circuit cause a corresponding change in electron flow in the collector circuit. Since the current in the emitter circuit consists of a d-c component owing to the bias battery plus an a-c component caused by the signal voltage, it may be inferred that the a-c component returns via the collector circuit. This concept is not strictly correct although the a-c components in the emitter and collector circuits are of the same order of magnitude. With suitable circuit voltages the collector current may be three times as large as the emitter current. This difference arises from a mutual action which takes place in the crystal surface surrounding the contact points. The apparent action is that the electrons flowing to the emitter alter the resistance of the crystal surface (semiconductor) and permit more electrons to flow from the collector electrode to the base electrode. Another helpful interpretation is that the negative potential of the

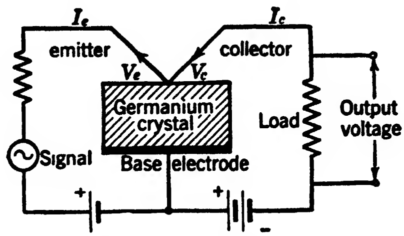


FIG. 6. Circuit of a Transistor used as an amplifier.

collector electrode lowers the resistance at the emitter contact and causes the signal voltage to be more effective.

The d-c characteristics of an experimental Transistor are shown in Fig. 7. There are four variables, two currents, and two voltages, with a functional relation between them. If two are specified, the other two are determined. In Fig. 7, the emitter and collector currents I_e and I_c are taken as the independent variables, and the corresponding

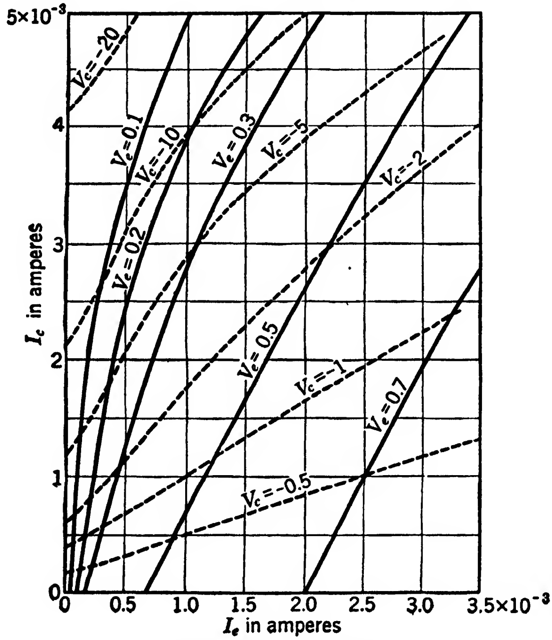


Fig. 7. Characteristics of a Transistor. (Courtesy Bell Telephone Laboratories.)

voltages V_e and V_c, measured relative to the base electrode, as the dependent variables.

Amplification of the signal voltage is produced in the Transistor amplifier because of the higher battery voltage and the high magnitude of the load resistor. Whereas the magnitude of the current changes in the emitter and collector circuits are of the same order, a much larger output voltage drop occurs across the load resistor. The average gain in the Transistor amplifier is about 20 db.

In comparison with the vacuum triode of equal capacity, (1) the Transistor does not require any power for heating the cathode, (2) it has a simple sturdy construction, (3) it appears likely that it will have a life of several thousands of hours, and (4) it uses a small amount of power. The Transistor consumes about 0.1 watt from bias sources

and delivers 25 milliwatts of useful output, thus having an over-all efficiency of 25 per cent.

In the early stage of its development the Transistor has some limitations. Its power output is limited to about 25 milliwatts per unit, or 50 milliwatts from a push-pull stage. Its noise level is higher than that of the vacuum tube, though this factor may not be objectionable for applications where high fidelity is not necessary. The low level of impedance on its input side and the reversed polarity of its input and output will require a new engineering of circuits in order to match its impedances to other circuits.

The properties of the Transistor indicate that it will ultimately replace electron tubes in a number of applications. Its light weight and low battery drain indicate its desirability for hearing aids and small portable radios. Again, its long life and low power requirements suggest its use in telephone and television repeater service for frequencies up to 5 or 10 megacycles.

Metallic Rectifier Cell.* Several combinations of substances have been discovered which when placed in the form of a sandwich show the property of unilateral conductivity. These sandwich combinations usually consist of a metal, a blocking layer, and some substance which is a semiconductor assembled as shown in Fig. 8. The blocking layer or rectifying layer exists at the surface junction of the metal and semiconductor and is formed by an electrochemical action or chemical action depending on the process for producing the sandwich. Some form of metal electrode is placed in contact with the semiconductor to carry away the current.

This form of rectifying unit conducts electrons fairly readily from the metal to the semiconductor but offers a high resistance to electron movement in the reverse direction. Actually the unit does conduct a little current in the reverse direction and the ratio of the current in forward or conducting direction to the reverse current is called the rectification ratio. It is easy to remember the direction of the electron movement by likening the metal electrode with its plentiful supply of free electrons to the space just outside the thermionic cathode and the semiconductor to the positive anode with a deficiency of available electrons. The theory of the action taking place in the metallic rectifier cell is not fully understood. Curiously, the movement of electrons from metal to semiconductor is opposite to that in the crystal rectifier discussed on the preceding pages. In considering this strange phenomenon, the physical differences in the two kinds of rectifiers should be noted. Thus, (1) the crystal rectifier has a small area of

* Also known as barrier cell and as a blocking-layer cell.

contact, whereas the metallic unit has a large one; (2) the crystal unit has a simple boundary or physical contact, whereas the metallic rectifier has a chemical union where the metal and semiconductor have a merging of molecules; and (3) the crystal rectifier has a thick semiconductor unit, whereas the metallic rectifier has a very thin layer. Apparently a polarized film or layer is built up between the metal and the semiconductor by the chemical or electrolytic action which takes place when the sandwich is formed. This layer permits part of the plentiful supply of free electrons in the metal to move across the thin semiconductor when it is positive, but it permits only a few to move under a reversed electric field.

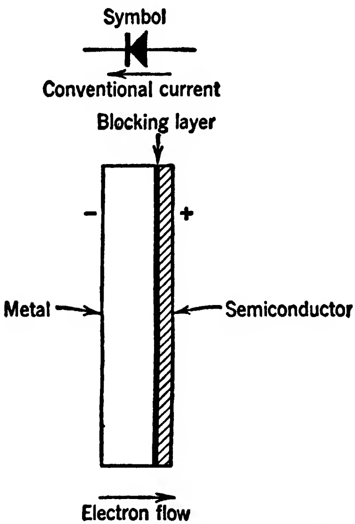


FIG. 8. Elements of a blocking-layer metallic rectifier and its symbol.

The *copper oxide rectifier* is a sandwich consisting of a layer of cuprous oxide on copper. This type of metallic rectifier was invented by Dr. Grondahl and its development followed a discovery made in 1920 while he was studying the photoelectric properties of cuprous oxide. Commercial forms of the copper oxide rectifier cells are shown in Fig. 9. These cells are produced by heating a copper disk or plate in a furnace to a temperature of approximately 1000 degrees F and then quenching it in water. This treatment produces a thin film of cuprous oxide with an outer layer of cupric oxide. The cupric oxide is then removed, leaving a thin layer (about 3 mils or 1000 molecules thick) of cuprous oxide on one side. Contact with the copper oxide surface can be made by holding a lead disk against the oxide surface under pressure or by electroplating a nickel coating on the surface of the oxide. The plated surface is preferred because there is a tendency for the lead disk to "flow" with time and to introduce an undesirable contact resistance in the circuit.

The unilateral characteristic of this cell is shown in Fig. 10 which gives the amperes in the forward direction of current and the milliamperes for the reverse or leakage current. The resistance of the unit is shown in Fig. 11 for a given temperature under a variation of voltage

from negative to positive. The resistance also varies with the current in the forward direction, being approximately inversely proportional to the current, as illustrated in Fig. 12. The resistance of the unit has a

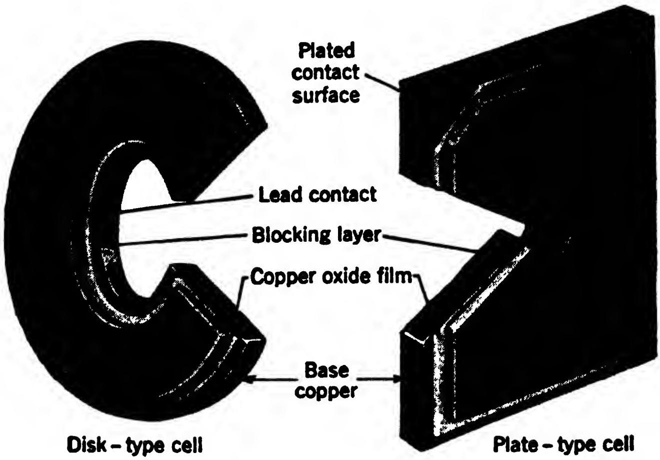


FIG. 9. Cross section of a copper oxide rectifier disk and plate-type cells. (Courtesy General Electric Company.)

negative coefficient of temperature which gives a variation of the current flow in the forward and reverse direction as shown in Fig. 13.

This rectifier is essentially a resistance device. Accordingly, the current squared times the resistance gives the power loss for both

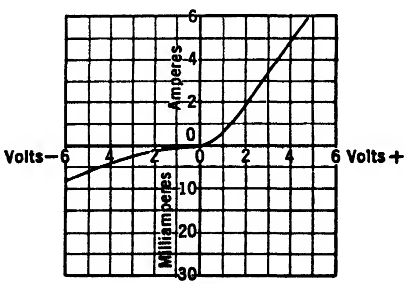


FIG. 10. Current-voltage characteristic curve of a copper oxide rectifier element. (Courtesy General Electric Company.)

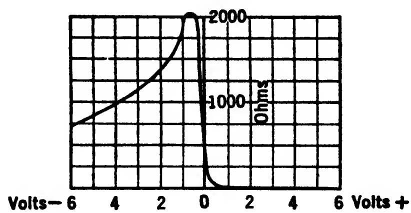


FIG. 11. Resistance-voltage characteristic curve of a copper oxide rectifier element. (Courtesy General Electric Company.)

forward and reverse current flow. These power losses raise the temperature and limit the rectified output of the device.

The various voltage, current, resistance, and temperature characteristics depicted in the figures for the copper oxide unit suggest certain

limitations which must be placed on the design and use of the copper oxide rectifier. In order to prevent an excess of reverse current, which represents a power loss and reduction in efficiency, the voltage impressed

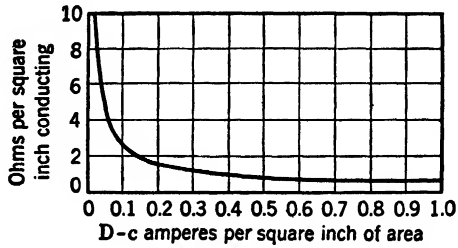


FIG. 12. Resistance-current characteristic of a copper oxide cell at 25 degrees C. (Courtesy General Electric Company.)

across a single cell must be restricted to from 8 to 11.5 volts. Also, in order to limit the leakage current and to prevent ultimate overheating, the maximum temperature for operation of the unit should be limited to about 45 degrees C. The output of the copper oxide rectifier may be increased by improved cooling which can be brought about by the addition of cooling fins and still more effectively by forced air cooling.

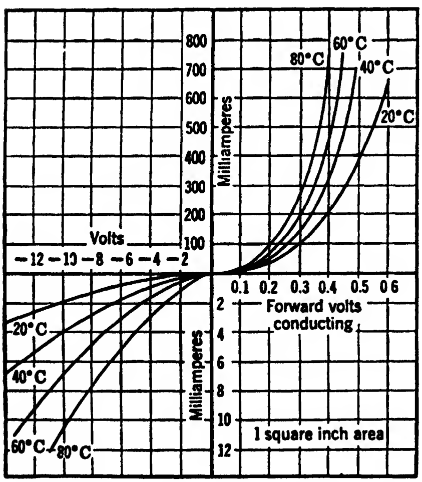
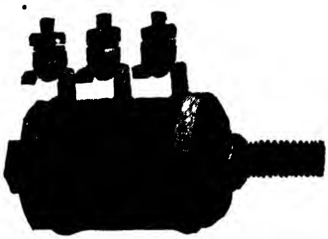


FIG. 13. Current-voltage characteristics of copper oxide rectifier cell. (Courtesy General Electric Company.)

The resistance of the copper oxide element increases with age. This "aging" effect on resistance increases with use and with rise in temperature. Fortunately, it appears that the resistance increases to a maximum stable value which is about twice the initial resistance. The result of aging is to reduce the efficiency of the rectifier with time and to change the value of the rectified voltage under load (regulation). Aging seems to be due to some inherent chemical or physical change in the rectifier element for which no remedy has been discovered to date. Commercial rectifier units carry a rating in which the effects of aging have been discounted so that the operation when the unit is new will be much better than the rating and guarantees.

Commercial copper oxide rectifiers usually consist of a number of elements in series or parallel combination. The number of elements in series depends on the applied a-c voltage and the number of elements placed in parallel depends on the required output current. Some copper oxide rectifier assemblies are shown in Fig. 14.



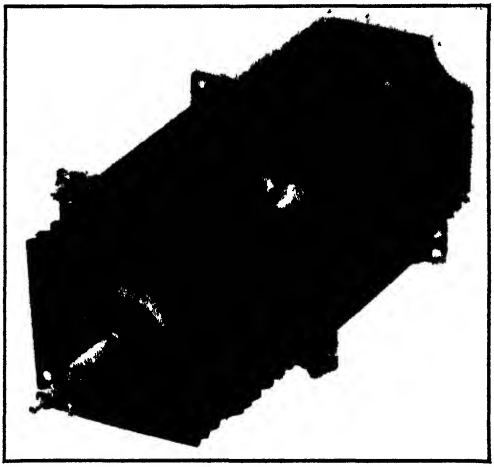


FIG. 14. Copper oxide rectifiers. (Courtesy General Electric Company.)

Copper oxide rectifiers are used in the electric power field for charging batteries, electroplating, motion-picture arcs, corrosion protection, ignitron exciter circuits, and similar applications. Their unidirectional conductivity makes them useful in series with d-c relays to guard against reversed polarity and in many circuits to permit current flow in one direction only. Copper oxide rectifiers are used also to permit measurement of alternating current with a d-c meter element. For this purpose four small copper oxide disks about $\frac{1}{4}$ inch in diameter are connected in the bridge circuit of Fig. 21, center, with a d-c meter in place of load. A similar grouping of four copper oxide disks sealed in a small metal container is known as a "varistor." Varistors are used in

large quantities for modulation and demodulation on long distance telephone carrier circuits (see Fig. 20, page 242).

The *selenium rectifier* is a blocking-layer type of rectifier which uses selenium for the semiconductor and a sprayed metal for a counter electrode. The principle of this device was discovered in the year 1883 though commercial application was not made in Europe until 1930 and in America a few years later. The selenium rectifying cell is illustrated in Fig. 15. It consists of a back plate of either iron or aluminum covered by a thin film of selenium. The film is given a series of controlled heat treatments to reduce the selenium to a suitable crystalline form. Next, a low melting alloy is sprayed onto the selenium surface. This layer is called the counter electrode and constitutes the active metal electrode of the sandwich unit. Subsequent electrochemical processes produced by the application of alternating current to the cell form a film or blocking layer between the counter electrode and the selenium.

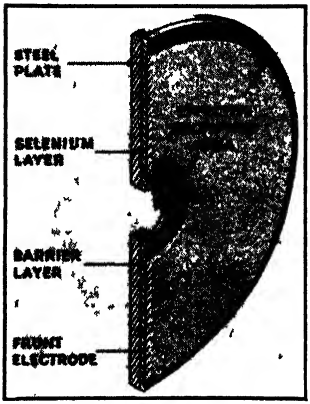


FIG. 15. Selenium rectifier cell. (Courtesy International Telephone and Telegraph Company.)

The current-voltage characteristic of a selenium cell per square inch of contact surface is shown in Fig. 16. The nature of the unilateral conduction is the same as that of a copper oxide unit with the electrons moving from the counter electrode into the selenium. The selenium cell has the same general characteristics as the copper oxide unit such as negative temperature coefficient of resistance, inverse relationship between resistance and current, and increase of resistance of the unit with age. However, it does differ from the copper oxide unit in some respects. It will withstand a maximum peak voltage of 25 volts per cell without damage. This means that for the same voltage output only one-half to one-third as many cells need be placed in series. Second, the selenium cell may be operated safely up to a maximum temperature of 75 degrees C. However, the continuous current output must be reduced as the temperature rises, so the advantage of higher temperature is not so important. The weight of the selenium rectifier for a given rating may be only one-seventh that for copper oxide for steel back plates or one-fourteenth when aluminum back plates are

used. One disadvantage of the selenium unit is that the resistance of the blocking layer is lower for direct current than for alternating current and that the blocking-layer resistance deteriorates when the cell is not in use though it is restored in a few seconds with the application of an a-c voltage. Hence the cell is not suitable for operating relays for reversed polarity on d-c circuits. Otherwise, the applications of the selenium rectifier are the same as those for the copper oxide type. The parts and assembly of a selenium rectifier unit are shown in Fig. 17.

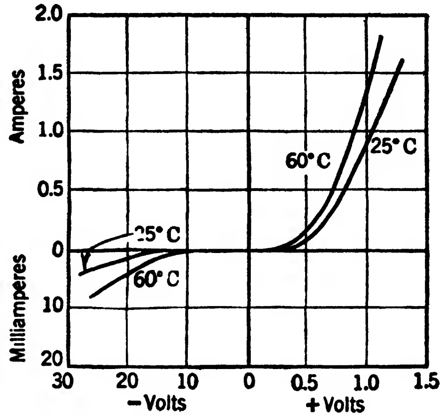


FIG. 16. D-c voltage characteristic of a selenium rectifier cell. (Courtesy General Electric Company.)

In 1946 new miniature selenium rectifier units were placed on the market which perform the same function as the 35 and 117 series of rectifier tubes. These new selenium units are more rugged and have a longer life than the rectifier tubes which they are designed to replace. One of these new units is illustrated in Fig. 18.

The *magnesium-copper sulphide rectifier* was invented in Austria and is manufactured in the United States under basic patents issued to Samuel Ruben in 1925. This blocking-layer metallic rectifier uses a magnesium plate for the metallic electrode and copper sulphide for the semiconductor. A back plate or electrode is placed in contact with the copper sulphide for collecting the current. The three pieces are assembled on a stud with radiator plates and held under pressure as illustrated in Fig. 19. The active blocking layer is formed between the magnesium and the copper sulphide by an electrochemical process. The forward movement of the electron flow is from the metal magnesium to the semiconductor copper sulphide. This rectifier unit must be protected from absorption of moisture to insure a long life; this necessity constituted one of the problems in the early manufacture and use of this rectifier. This rectifier cell has a rectification ratio of 75 to 1, which is lower than that of other disk rectifiers and tends toward a lower rectifier efficiency.

The advantages claimed for the magnesium-copper sulphide rectifier are (1) its ability to withstand heavy current overloads, (2) an

extreme range of operating temperature, plus (3) a high normal current density. This unit has a self-healing rectifying film which will

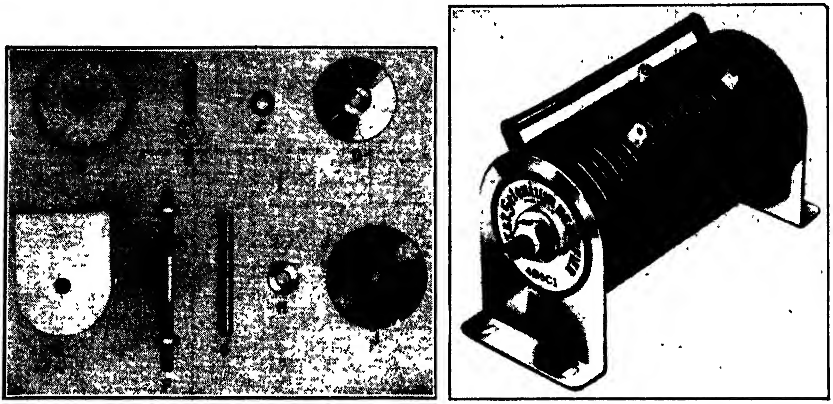


FIG. 17. Parts of a selenium rectifier and an assembled selenium rectifier. (Courtesy International Telephone and Telegraph Company.)

seal a puncture due to overload or a transient voltage surge. This rectifier will operate under a range of temperature from −70 to 135 C and has been tested in the laboratory under loads that produced temperatures as high as 275 degrees C. The normal current density depends on the method of cooling employed and lies in the range of 25 to 50 amperes per square inch.

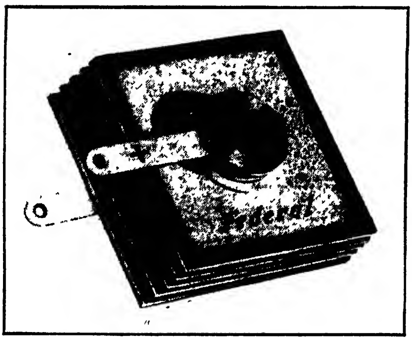


FIG. 18. Miniature selenium rectifier. (Courtesy International Telephone and Telegraph Company.)

The magnesium-copper sulphide rectifier may be employed for the same applications as the metallic rectifiers previously discussed. A commercial rectifier of this type is shown in Fig. 20.

Summary on Metallic Rectifiers. Metallic rectifier cells are resistive units, and hence provide a unity power factor load as far as the direct rectification process is concerned. These cells can be used in series and parallel combination for any type of rectifier circuit. Three typical circuits for metallic rectifiers are shown in Fig. 21. Several other circuits will be given in Chapter XIV.

The design of a metallic rectifier for a given application will be determined by the load amperes, the desired voltage regulation, and the

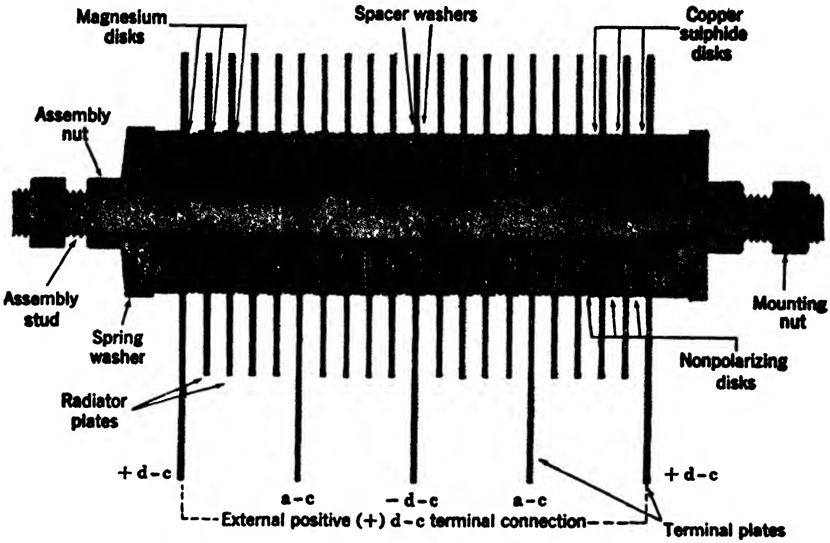


FIG. 19. Construction of a magnesium-copper sulphide rectifier. (Courtesy P. R. Mallory & Company, Inc.)

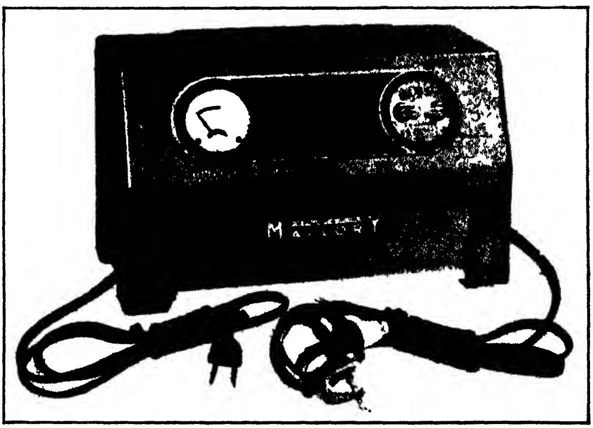


FIG. 20. Commercial battery-charging unit using magnesium-copper sulphide rectifying disks. (Courtesy P. R. Mallory & Company, Inc.)

characteristics of the type of rectifier cell employed. The important cell characteristics are (1) the permissible inverse peak voltage, (2) the permissible temperature rise, and (3) the permissible current density. All metallic rectifier cells have negative coefficients of resistance

though the resistance is approximately constant in the normal current density range. If the resistance is considered constant, the voltage regulation curve would show a linear fall of potential with load. The decreasing resistance characteristic with load (Fig. 12) causes the regulation to be better than linear, as shown in Fig. 22. To provide satisfactory regulation the total resistance of the rectifier cells at full d-c

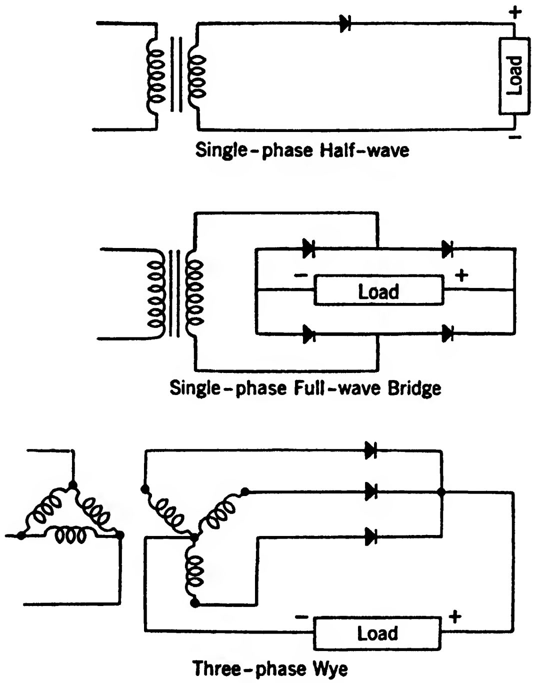


FIG. 21. Typical circuits for metallic rectifiers.

load should not exceed 10 to 15 per cent of the load circuit resistance. The number of rectifier cells required is determined by the permissible inverse peak voltage and the d-c no-load voltage. Lastly, the size (and number in parallel) of the rectifier cells is determined by the d-c load current. The manufacturers use empirical methods for designing rectifiers for applications. The student may apply the preceding statements of this paragraph, the data on Figs. 12, 13, and 16, plus the data of Table 1 to make an approximate design or to compute the probable rating of a rectifier having an unknown rating.

The efficiency of a metallic rectifier is the ratio in per cent of the d-c power output to the total power input where both output and input are measured with a-c wattmeters. The efficiency of metallic rectifiers

is reduced by the reverse current flow as well as by the I^2R loss within the rectifier units. Efficiency of rectifiers will be given further consideration in Chapter XIV. Since the term "efficiency" as applied to metallic rectifiers is subject to different interpretations, the manufacturers of metallic rectifiers prefer to use the term "conversion ratio." *The conversion ratio of a metallic rectifier is the ratio, in per cent, of the product of the average values of d-c voltage and d-c current output to the total a-c power input.*

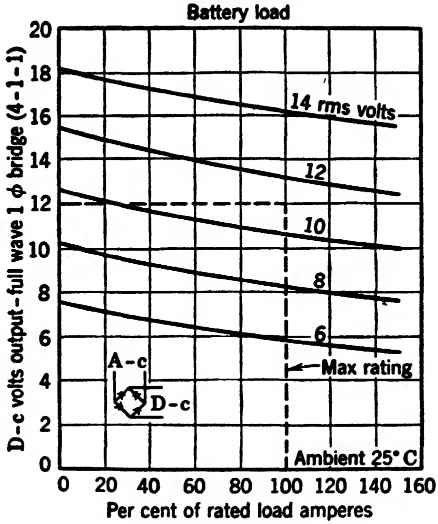


FIG. 22. Voltage-regulation curves of a selenium rectifier. (Courtesy General Electric Company.)

A summary and comparison of the characteristics of copper oxide, selenium, and magnesium-copper sulphide rectifiers is given in Table 1. The student should study this table.

TABLE 1.* COMPARISON OF METALLIC RECTIFIERS

Type of Metallic Rectifier Cell	Peak Inverse Volts per Cell		Current Density Amperes per Square Inch (full-wave rectification)		Maximum Recommended Temperature (degrees)		Efficiency (over-all per cent)	
	Normal operation	Intermittent service or reduced life	Conventional cooling	Forced draft cooling	Normal operation	Intermittent service or reduced life	For 1 ϕ	For 3 ϕ
Copper oxide	8.75–11.5	30.0	0.3	0.5	35 C	45 C	60–70	75–80
Selenium	25.0	54.0	0.3	0.75	35 C	75 C	60–70	70–75
Magnesium-copper sulphide	5.0	5.0	25.0	50.0	40 C	85–130 C	37–45	52.7

* Several factors enter into the rating and operation of metallic rectifiers so that the values suggested in this table should be considered as approximate.

Problems

In the solution of the following problems, assume that the peak applied voltage is 1.41 times the applied rms alternating voltage and that the d-c average voltage is 0.318 times the peak for half-wave rectification and 0.636 times peak for full-wave.

1. An a-c voltage of 60 volts (rms) is to be rectified by copper oxide cells to deliver 2.0 amperes of direct current with full-wave rectification under normal operation and with convectional cooling. Calculate (a) the number of cells, (b) the area of the cells, and (c) the average d-c voltage under load.

2. An a-c supply of 120 volts rms is applied to selenium rectifier stacks (full-wave) to give 100 milliamperes direct current. How many cells will be required per stack and what should be their size for intermittent service? What will be the d-c average voltage? What should be the magnitude of the load resistance? What will be the resistance drop across the rectifier?

3. A metallic rectifier is to be used near a furnace where the ambient temperature may rise to 100 degrees C. What type of cell should be used?

4. Using Fig. 22, calculate the per cent voltage regulation of a selenium rectifier which charges a 6-volt storage battery (with taper charge) up to a near zero rate with a final voltage of 7.5 volts.

5. Repeat Problem 4 for a 12-volt battery, having a final (zero charging current) voltage of 15.5 volts.

REFERENCES

Hamann, C. E., and E. A. Harty, "Fundamental Characteristics and Applications of Copper-Oxide Rectifiers," *Gen. Elec. Rev.*, August 1933.

Clark, Carole A., "Selenium Rectifier Characteristics, Applications, and Design Factors," *Elec. Commun.*, **20**, 1941.

Cornelius, E. C., "Germanium Crystal Diodes," *Electronics*, February 1946, p. 118.

Chapter XIV

RECTIFICATION AND INVERSION

Rectification of electric current is the process of changing an alternating current into a unidirectional current. A unidirectional rectified current is pulsating in character and it is generally necessary to smooth out the pulses by means of a filter in order to utilize the resultant current and voltage. Rectification of alternating current is exceedingly important because (1) it is more simple and economical to generate and distribute electric energy as alternating current, and (2) many applications of electric energy in the field of electric power and communication require the use of a unidirectional or direct current. Some of these applications requiring the use of direct current are the power supply to the anodes of electron tubes, electroplating, chemical processes, charging of storage batteries, and operating series railway motors and adjustable-speed d-c motors.

Rectification of alternating current or the conversion of alternating current to direct current can be brought about by (1) electronic methods and (2) commutating methods. This textbook is concerned primarily with the electronic methods, but for purposes of comparison some of the commutating methods will be given brief consideration. Inversion, the process of converting a direct current to an alternating current, may likewise be performed by both electronic and commutating means.

Commutating Rectifiers. Commutating rectifiers may be of the vibrating contact, the synchronous motor-driven contact, or the rotary type of commutator. A *vibrating rectifier* consists of an electromagnet which vibrates a spring containing movable contacts. The period of vibration is controlled or tuned by the pulses or alternations of the current, and the movable contacts serve to reverse the direction of current flow so that the outgoing current is rectified or unidirectional. Devices employing this principle were used in earlier years for charging storage batteries and for rectification in telephone-ringing machines. Today the commutating vibrator is widely used as

an inverter for changing direct current to alternating current in automobile and other portable radio sets, and in some sets it performs the double function of inversion followed by rectification.

The circuit for a typical vibrating type of inverter is shown in Fig. 1. The closing of switch S energizes the driving coil over a path from the ground (+ side of the battery), the driving coil, and the right half of primary winding P to the negative side of the battery. The resulting current energizes the driving magnet and pulls the vibrator V to the

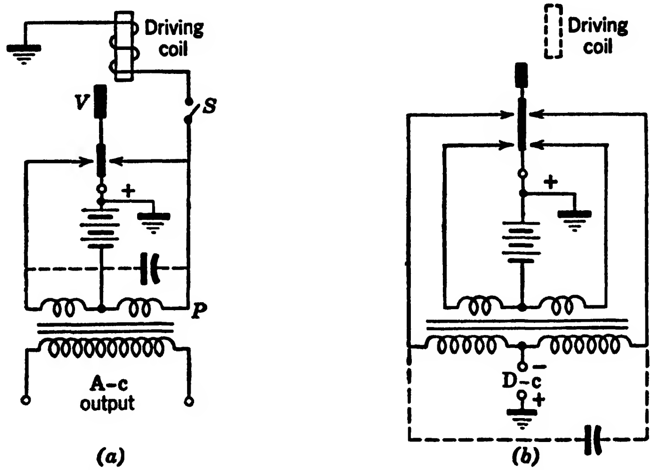


Fig. 1. Vibrator types of inverters and rectifiers: (a) inversion; (b) combined inversion and rectification.

right-hand contact. The closing of this contact (1) sends a relatively large d-c pulse of current through the right half of winding P and (2) places a short circuit around the driving coil. The first action induces a voltage in the output side of the transformer, and the second serves to de-energize the driving magnet. The spring of the vibrator now swings the movable member V to the left until it engages the left-hand contact where it sends a pulse of current through the left half of the primary winding P. This new current pulse reverses the flux in the transformer core and reverses the voltage induced in the output side of the transformer. In the next instant the vibrator V swings again to the right owing to its spring action plus the reestablished current through the driving coil. The repetition of the foregoing action produces an alternating voltage at the output terminals of the step-up transformer. This a-c output voltage may be rectified by a vacuum diode, by a small metallic rectifier, or by the addition of two contacts at the

vibrator and a mid-tapped secondary to the inverter circuit. The additions are shown in part *b* of Fig. 1. There the driving circuit (not shown) acts in the same manner as before. A study of this circuit will show that the added contacts serve to reverse the voltage between the mid-tap of the output winding of the transformer and ground in such a way as to make the resulting voltage pulses unidirectional, and thus give full-wave rectification. A smoothing filter must be added to give an output suitable for the power supply for radio receiving sets.

The preceding explanation has been an over-simplification of the complete theory of action involved in the vibrator type of inverter and combined inverter and rectifier. Condensers are used in the circuits to reduce arcing at the contacts but the complete theory of action and design involves a proper balance and tuning of the L and C constants in the circuits.*

A new type of mechanical rectifier or *contact converter* was developed by Siemens-Schuckert in Berlin during World War II. This contact converter is capable of rectifying currents as high as 10,000 amperes at 400 volts with efficiencies of the order of 98 to 99 per cent. The principle of rectifying alternating current by reversing the circuit in step with the alternations is old in the art, but for sizable currents no satisfactory process has been developed until recently. In theory it is necessary only to reverse the circuit while the voltage and current pass through the zero point on the cycle. To effect this requirement has been impractical because of the short time available for mechanically moving the contacts at the time when the rate of current change is the maximum. Any phase-angle difference between the voltage and the current naturally complicated the problem of commutation. If the wave form of the circuit can be modified to hold the voltage at near zero for a short time interval at the points of voltage and current reversal, it is obvious that the mechanical problem of reversal would be simplified. This important advantage is brought about in the new contact rectifier by inserting saturable inductances or chokes in the alternating-current leads. These chokes saturate at relatively low values of flux and current, so that they produce a high impedance to current change at low values of current (region when passing through zero) and little impedance at higher current values. These chokes act in a manner similar to the peaking transformer although their function is much different. The application of the saturable choke gives the necessary delay in the voltage and current change and permits the mechanical reversal of contacts operated by a small

* For more complete details the reader is referred to the *Vibrator Data Book* published by P. R. Mallory & Co., Inc., copyright 1947.

synchronous motor. Many mechanical and electrical refinements are necessary to make the contact converter successful. The discussion of these refinements is out of the province of this book. Contact converters are now being manufactured in the United States.*

Rotary commutating rectifiers have been of three types: simple commutators, motor-generator sets, and rotary converters. The simple commutator device consists of a synchronous motor driving a commutator at such a speed that the polarity is reversed at the proper rate to give a rectified or unidirectional current in the output. This device was widely used in early radio transmitters and X-ray equipment and today has considerable application for rectifying high voltages for precipitation processes in the removal of smoke, dust, and chemical byproducts. The motor-generator rectifier consists of an a-c motor operating from an a-c system which drives a d-c generator. This process is an electrical-mechanical plus a mechanical-electrical conversion process. It has the advantage of good control of the d-c output voltage and the disadvantages of high initial cost, lower efficiency, and the complications of rotating equipment. This system is being replaced by electronic methods of conversion. The rotary converter consists of a single rotating unit which consumes alternating current in motor action and gives forth a d-c output. The output is partly the result of pure rectification and partly of a conversion process. This device has a lower cost and a much higher efficiency than the motor-generator device. Its disadvantage lies in inflexibility in the control of the output (d-c) voltage. This unit once had a wide application in the electric railway field but is being replaced by electronic rectifiers.

All these commutating devices used for conversion of alternating current to direct current may be employed equally well through a reverse process to bring about inversion.

Electrolytic Rectifiers. An electrolytic rectifier consists of two dissimilar electrodes placed in an electrolyte. One type of cell uses an aluminum plate and a lead plate for the electrodes and a solution of ammonium phosphate for the electrolyte. In this cell electrons pass readily from the aluminum electrode to the lead electrode but not vice versa. Thus the cell may be used as a single-wave rectifier. Obviously, the phenomenon is of electronic origin and the action is explained by the presence of a barrier at the surface of one electrode which has a high resistance to the passage of electrons from the electrolyte into the heart of the electrode but little resistance to a movement in the reverse

* For more information the reader is referred to the article, "A Mechanical Rectifier" in the April 1948 issue of *Power Generation.*

direction. The work function of the material at the surface of the electrodes determines the resultant action.

The electrolytic rectifier was widely used for charging storage batteries during the pioneering days of radio. It had the advantage of simplicity and low cost and the disadvantage of low efficiency and short life. Its low efficiency was due to a rather high internal voltage drop plus a loss from a small reverse current flow. The electrolytic rectifier has been obsolete since 1930.

Single-Phase Rectifier Circuits. The simplest circuit for rectifying single-phase alternating current gives half-wave rectification. Such a circuit is shown in Fig. 2, part *a*. This circuit can be simplified by reducing it to the schematic or equivalent form of part *b* in the same figure. Here a new symbol represents the rectifying unit and *the arrow of the symbol indicates the direction of conventional current flow.* When considering this or any other rectifying circuit the student may wish to know the magnitude of the d-c voltage and current, the regulation of the load voltage, and the efficiency to be expected from the rectifying process. All these values depend upon a number of variables such as the transformer constants and characteristics, the characteristics of the rectifying unit, and the nature of the load. These variables may make an analytical solution of the problem rather difficult. However, it is possible to make certain assumptions which reduce the circuit to an ideal basis from which a helpful analysis may be made. Referring to Fig. 2, part *a*, assume (1) an ideal transformer without resistance, without core loss, and with zero regulation, (2) that the impressed emf is a pure sine wave, (3) that the rectifying unit has zero resistance in the forward direction of electron movement and infinite resistance in the reverse direction, and (4) that the load is a pure ohmic resistance. With these assumptions, let an alternating voltage of effective value E_{rms} be impressed on the left side of Fig. 2, part *b*. The impressed voltage E_{rms} may be represented by the sine wave of part *c* of Fig. 2 and this voltage will produce a rectified half-wave of current as shown in part *d* of the same figure. This half-wave of current is of a sine form since $i = (E_m/R_L) \sin \omega t$. During the last half of the cycle the rectifier will block the electrons and no current will result. The same loop or half-wave of current will flow on both the input and output sides of the rectifier. This current flowing through the load resistance R_L will produce an iR_L voltage drop which is likewise of a sine form. The d-c values of pulsating wave forms are the average values. By the methods of calculus, the area of the rectified sine wave form of current in Fig. 2*d* is the integral from 0 to π of $i\,d\theta$

which equals $2I_{max}$. The average value of this area spread over a complete cycle is

$$I_{avg} = I_{dc} = \frac{2I_m}{2\pi} = 0.318I_{max} \tag{1}$$

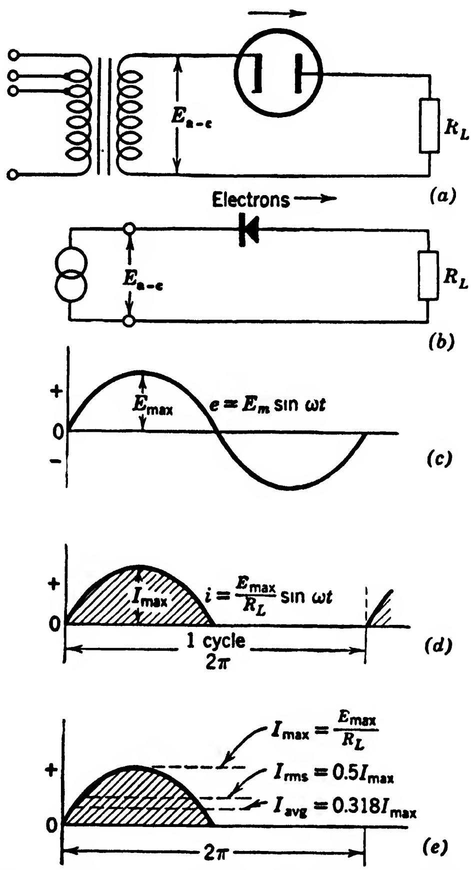


FIG. 2. Circuit and waves for half-wave rectification.

The instantaneous voltage equals iR_L, and the same deductive process shows the average voltage on the load side of the filter for the complete cycle to be

$$E_{avg} = E_{dc} = 0.318E_{max} = 0.45E_{rms} \tag{2}$$

The effective value of the current on the input side is the root-mean-square value. Obviously, the effective value of the first half-cycle of current is $E_{max}/\sqrt{2}$. The mean square of this is $E_{max}{}^2/2$. Now if this

latter value is spread over the complete rectifying cycle, its mean square value is

$$\frac{1}{2} \times \frac{I_{\max}{}^2}{2} = \frac{I_m{}^2}{4}$$

and the root-mean-square of the latter is

$$\frac{I_{\max}}{2} = 0.5 I_{\max} \tag{3}$$

These values are indicated in part *e* of Fig. 2. Summarizing,

Input Side	Output Side
E_{rms} = effective impressed voltage	$E_{\text{dc}} = 0.45 E_{\text{rms}}$
$I_{\text{rms}} = 0.5 I_{\max}$	$I_{\text{dc}} = 0.318 I_{\max}$

The heating losses (I^2R) which occur in the various parts of the complete rectifier circuit are increased by the irregular wave forms of current which are inherent in the rectifying process. This may be illustrated by comparing the heating loss produced in the load resistance R_L by a smooth or filtered d-c current with that resulting from the current wave of Fig. 2. The resulting ratio may be termed the ratio of rectification. Thus from the preceding table

$$\underset{\text{(half-wave rectification)}}{\text{Ratio of rectification}} = \frac{(I_{\text{dc}})^2 R_L}{(I_{\text{rms}})^2 R_L} \times 100 = \frac{(0.318 I_m)^2}{(0.5 I_m)^2} \times 100$$

$$= 40.6\% \tag{4}$$

The ratio of equation 4 is sometimes called the efficiency of rectification. The latter terminology is somewhat misleading since the over-all power efficiency for the assumed conditions (zero losses) must be 100 per cent. The real significance of the ratio of rectification is that it gives a qualitative indication of the increased heat losses that occur wherever a nonsinusoidal and pulsating current flows through resistance elements in the actual rectifier circuit such as the transformer windings and in the load without a filter. A second method of expressing the increased heat losses is by the term current *form factor* which is the ratio of root-mean-square to average value.

The capacity of the input transformer to a rectifier is used inefficiently when supplying power for half-wave rectification because only one-half of the sine wave of current is passed and the secondary winding carries a d-c component of current which magnetizes the iron core and increases the core losses. These factors plus the increased I^2R losses suggested in the preceding paragraph result in a reduction of the

permissible transformer output under rectifier loads. This reduction is sometimes expressed by the term utility factor. *Utility factor* is

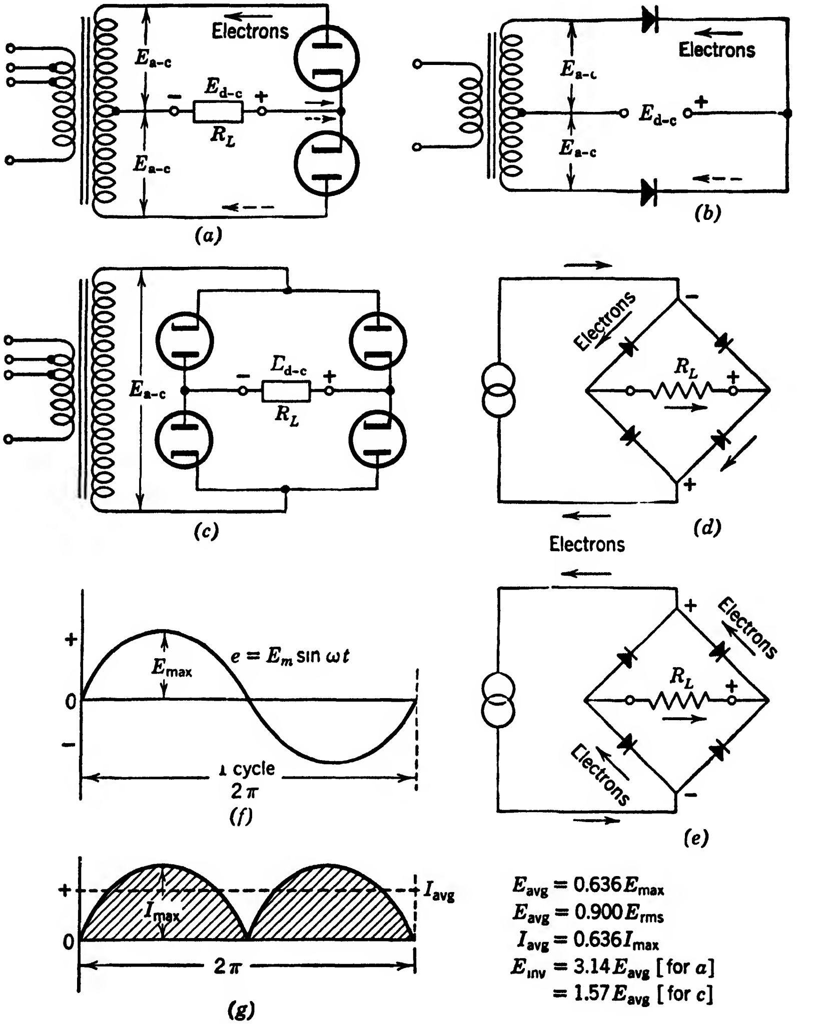


FIG. 3. Circuit and waves for full-wave rectification. (Conventional current flows in direction of boldface arrows.)

the ratio of the permissible rectifier load without overheating to the normal unity power factor load that the transformer will carry.

There are two types of circuits used for full-wave single-phase rectification. One of these circuits using a transformer with a mid-tap in the secondary winding is shown in parts *a* and *b* of Fig. 3. For one

half-cycle of the impressed a-c voltage E_{rms} the upper half of the transformer secondary and the upper diode conducts current, and for the second half-cycle the lower half of the transformer and the lower diode operates. If a pure sine wave of voltage like part *f* of Fig. 3 appears across each half of the transformer secondary winding, both halves of the cycle will be rectified as shown in part *g* of this figure. The connection to the midpoint on the secondary of the transformer serves to reverse the polarity on the lower rectifier with respect to the upper one. Since both half-waves of current pass through the transformer (though dividing in the secondary), there is no d-c component of flux in the transformer core to increase core losses, and the current in the primary is normal. The second type of circuit to give full-wave rectification is shown in parts *c, d,* and *e* of Fig. 3. This is known as a bridge circuit since it uses four rectifying units connected in a form similar to the Wheatstone bridge. The rectified currents flow through two diodes in series and the action for the two halves of each cycle is clearly shown in parts *d* and *e* of the figure. The resulting impressed voltage and current waves are the same as for the first full-wave rectifier circuit (Fig. 3, parts *f* and *g*).

For the same assumptions made for the preceding half-wave rectifier circuit, the relation between the input and output sides of either full-wave rectifier circuit may be readily calculated. Since both halves of the cycle are rectified, the current and voltage on the input side are normal effective values and those on the d-c or output side are average values. Thus

$$E_{rms} = \text{applied rms value} = \frac{E_{max}}{\sqrt{2}}$$

$$I_{rms} = \frac{E_{rms}}{R_L} = \frac{E_{max}}{\sqrt{2}R_L} = 0.707\,\frac{E_{max}}{R_L}$$

$$E_{dc} = E_{avg} = 0.636E_{max} = 0.9E_{rms}$$

$$I_{dc} = I_{avg} = 0.636I_{max} = 0.9I_{rms}$$

$$\underset{\text{(full-wave rectification)}}{\text{Ratio of rectification}} = \left(\frac{I_{dc}{}^2R_L}{I_{rms}{}^2R_L}\right) \times 100$$

$$= \left(\frac{0.636}{0.707}\right)^2 \times 100 = 81.2\% \qquad (5)$$

Equation 5 gives a concept of the extra I^2R losses occasioned by the wave form of the resulting currents.

The rectified voltage and current output of any rectifier consists of a series of unidirectional waves or ripples. For some applications these variations are not objectionable but for others they must be smoothed out by filters. For all cases the relative magnitude of the ripple is important in the comparison of rectifying circuits. The comparison is made in terms of ripple factor. *Ripple factor is the ratio of effective value of the alternating components of the rectified voltage or current to the average value.* In equation form ripple factor is

$$r_f = \frac{\text{effective rectified a-c component } (I_{\text{rms}}')}{\text{average current } (I_{\text{dc}})} \tag{6}$$

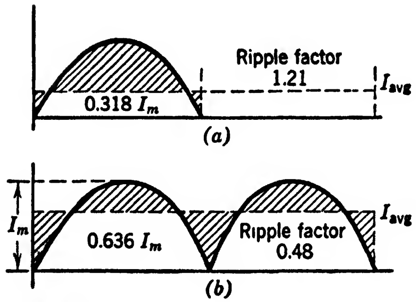


FIG. 4. Alternating component of rectified current: (*a*) half-wave; (*b*) full-wave.

The alternating component of current for half-wave and full-wave rectification using a sine wave is shown by the shaded portions on Fig. 4. The d-c or average component is represented by all the area underneath the dotted line for I_{avg}. The ripple factor for the half-wave rectification can be calculated from the preceding calculations by use of the following steps. The instantaneous a-c ripple component i' in Fig. 4 is

$$i' = i - I_{\text{dc}}$$

and the total rms value of the ripple component I_{rms}' by the calculus is

$$I_{\text{rms}}' = \sqrt{\frac{1}{2\pi}\int_0^{2\pi} (i - I_{\text{dc}})^2 \, d\theta}$$

$$= \sqrt{\frac{1}{2\pi}\int_0^{2\pi} (i^2 - 2I_{\text{dc}}i + I_{\text{dc}}^2) \, d\theta}$$

The first term in this expression is the rms value of the total current I_{rms}. That part of the second term $\int_0^{2\pi} i \, d\theta$ integrates to the average value I_{dc} as developed on page 333, and the last term is simply I_{dc}^2 after the limits are applied. Thus

$$I_{rms}' = \sqrt{I_{rms}^2 - 2I_{dc}^2 + I_{dc}^2} = \sqrt{I_{rms}^2 - I_{dc}^2}$$

and

$$r_f = \frac{\sqrt{I_{rms}^2 - I_{dc}^2}}{I_{dc}} = \sqrt{\left(\frac{I_{rms}}{I_{dc}}\right)^2 - 1} \tag{7}$$

Substitution of values from the preceding development into equation 7 gives

$$\text{Half wave } \frac{I_{rms}}{I_{dc}} = \frac{0.5I_{max}}{0.318I_{max}} = 1.57$$

$$r_f \text{ (half wave)} = \sqrt{1.57^2 - 1} = 1.21$$

$$\text{Full wave } \frac{I_{rms}}{I_{dc}} = \frac{0.707I_{max}}{0.636I_{max}} = 1.11$$

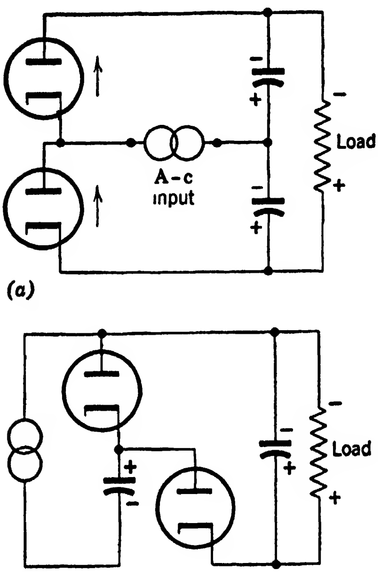
$$r_f \text{ (full wave)} = \sqrt{1.11^2 - 1} = 0.482$$

The d-c and a-c components of rectified current can be measured by inserting a d-c and an a-c ammeter (plus current transformer) in series in the circuit.

The preceding discussion of ripple factor and a further inspection of Fig. 4 will explain why the ratio of rectification is lower than might be expected. The rectified current may be considered as consisting of two component currents, one a steady direct current of the average value useful for d-c application and the other an a-c ripple current that contributes to the I^2R or heat loss in the load. This heat loss may constitute a sizable factor in reducing the efficiency of d-c motor loads operating on rectifiers. In half-wave rectification the a-c component is very large (1.21 times) compared to the d-c component.

A further comparison of the three single-phase rectifying circuits may be of interest. The half-wave rectifier circuit is simple, uses only one rectifier unit, has a high ripple factor, makes inefficient use of its transformers, and has a low ratio of rectification. The full-wave rectifier circuits have a smaller ripple factor and a much higher ratio of rectification. The circuit using a mid-tap transformer requires only two rectifying units but the transformer is special and more costly. Besides the necessity of a mid-tap, the secondary must have twice as many turns of somewhat smaller wire; hence a higher cost. In addition, the distorted wave form of the current in the two secondaries adds to the I^2R losses. Where diode rectifying tubes are used this circuit

is standard because tubes have a relatively high initial cost and maintenance cost. The full-wave bridge circuit uses a simple two-circuit transformer but four rectifying units. This circuit is standard for use with the blocking-layer rectifiers because the rectifier cells can be separated into two parts to comprise the two rectifying units for each path through the bridge circuit. This circuit uses a simple transformer with a single winding on the secondary. Under the ideal conditions originally assumed, there are no extra losses in this transformer arising from rectification. Under actual conditions of rectification there is some wave distortion resulting in a utility factor of less than one.



(a)

(b)

FIG. 5. Voltage doubler circuits: (a) full-wave; (b) half-wave.

The rectified output voltages of the single-phase circuits considered above are less than the input effective voltages. If an output voltage greater than the input is desired, it can be obtained by either of the voltage doubler circuits given in Fig. 5. In part *a* full-wave rectification is secured by connecting two diodes in two arms of a bridge circuit with condensers in the other arms. While rectifying one-half a cycle of an a-c voltage, the upper diode will charge the upper condenser to a potential approaching E_{max}. Similarly, on the second half-cycle, the lower diode will charge the lower condenser to a potential approaching E_{max}. Since the diodes rectify voltages that are series aiding, the total voltage across the load and the two condensers in series will approach the value $2E_{max}$. The actual voltage will depend upon the current drain (discharge) by the load and upon the cathode-anode voltage drop across the diodes. Part *b* of Fig. 5 shows another circuit for doubling the output voltage. Here on the first half-cycle of impressed voltage the diode on the left charges the condenser on the left to a potential approaching E_{max}. On the second half-cycle the peak voltage across the diode to the right is the impressed E_{max} plus the voltage due

to the charge on the left-hand condenser. The rectification of this combined voltage approaching $2E_{max}$ charges the condenser on the right to a load voltage approaching $2E_{max}$ as in the first circuit. The circuit of part *a* gives better regulation, higher ripple frequency, and lower voltage across the condensers; the circuit of part *b* has a common input and output terminal.

The actual resistance and power loss of a rectifying unit will alter the voltage and current relations and efficiencies developed in the preceding discussion where an ideal status was assumed. The rectifying units in any of the circuits considered may be high-vacuum diodes, gas or vapor diodes, or blocking-layer rectifiers. The differences in the characteristics of each of these units were discussed in earlier articles. It is possible to make some generalization covering these characteristics and thereby make a closer approximation for some factors of rectification. Thus for vacuum diodes and blocking-layer units the assumption may be made that the resistance is constant. If the resistances of such a rectifier are represented by the symbol R_R, leaving the load resistance R_L, then for any given set of conditions the rectified current and the rectified voltage will be reduced in the following ratio.

$$\underset{\text{(high-vacuum tubes)}}{\text{Ratio of reduction}} = \frac{R_L}{R_L + R_R} = \frac{1}{1 + (R_R/R_L)} \tag{8}$$

In gaseous and vapor tubes the arc drop or cathode-anode drop is practically constant. If this drop is represented by V_R, then the instantaneous values of the output voltage will be the impressed value minus V_R as shown in Fig. 6. Any general analytical solution for this case becomes somewhat involved and of doubtful value.

The output voltage E_{dc} of a rectifier using vacuum diodes or blocking-type cells will decrease with load by an amount determined by the reduction factor (equation 8). If it is necessary to hold the value of E_{dc} to a constant value, it is necessary to vary the input voltage E_{rms}. This can be accomplished in steps by changing taps on the primary of the input transformers in Figs. 2 or 3, or more exactly by some form of input voltage regulator.

Multiphase Rectifier Circuits. Multiphase rectifier circuits are used whenever moderate or large magnitudes of d-c power are applied. Such circuits have the advantage of utilizing the standard three-phase power distribution system and they furnish smoother rectified voltage and current waves at higher efficiencies. A simple three-phase circuit illus-

trating the principle of multiphase rectification is given in Fig. 7. Power is supplied through a three-phase, delta-wye connected transformer to a three-anode unit rectifier. The rectifying units may consist of vacuum diodes, gaseous diodes, phanotrons, ignitrons, or excitrons having all cathodes connected together, or a multielectrode tank rectifier having a single-cathode mercury pool may be used. The d-c load is connected to the common cathode terminal (+) and to the neutral of the transformer secondary. All half-wave multiphase circuits require that one load terminal arise from a neutral or common point in the transformer supply secondary. The effective a-c voltage for rectification E_{rms} is always the voltage measured from the transformer neutral to the anodes of the rectifier.

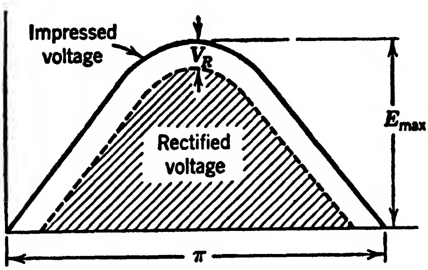


FIG. 6. Effect of a constant arc drop upon the output voltage of a rectifier.

An analysis of the action of multiphase rectification may be made by assuming ideal rectifying units with zero resistance, an ideal transformer, a pure resistance load, and a sine wave of applied voltage. The variation of voltage across the anodes during one cycle is depicted in the middle section of Fig. 7 and the resulting current waves in the lower section of the figure. The d-c voltage waves will have the same form as the current waves. The rectified waves have ripples per cycle equal to the number of anodes but the ripples are lower in magnitude giving a lower ripple factor and a much smoother wave. Under the ideal theoretical conditions assumed, the rectified current flows to the anode having the highest positive potential and that means that each of three anodes will carry the current for $2\pi/3$ part of each cycle. Since the ripples on the load side of the transformer are equal and continuous, the d-c or average value of the current or voltage will be the average value of a single ripple wave. The peak value of the ripple voltage is the maximum value of the impressed voltage E_{max}. As the number of phases and anodes increase, the number of ripples and the average d-c voltage increase and the voltage approaches the maximum value of E_{max}.

One method of obtaining a low ripple from a three-phase circuit and a standard transformer is shown in the circuit of Fig. 8. In practice, this circuit, using a delta-connected transformer, generally em-

ploys the blocking-layer rectifier. Six rectifier tubes of any type can be employed though the higher cost of the tubes may make their use uneconomical. The transformer phase voltages and the resulting current output with six ripples per cycle is shown at the right of Fig. 8.

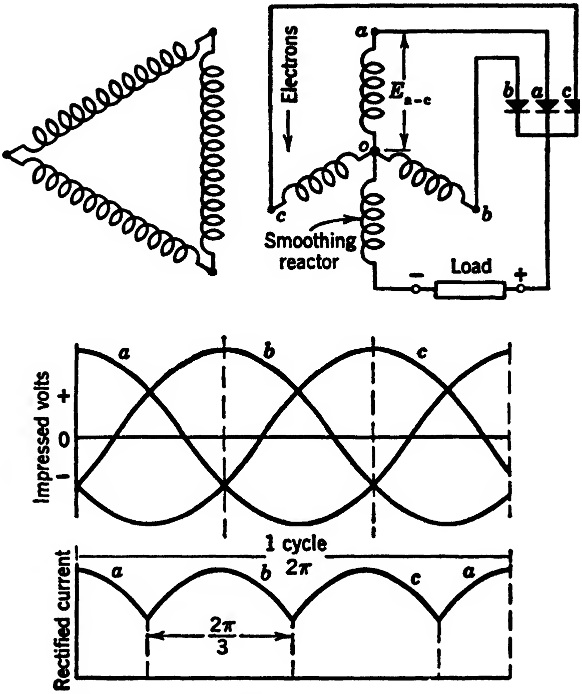


FIG. 7. Half-wave, three-phase rectifier circuit and voltage-current waves (boldface arrows show direction of conventional current).

Obviously, this three-phase circuit is analogous to the full-wave single-phase bridge type of circuit.

A schematic circuit for six-phase rectification using a special transformer and a rectifier having six anodes is given in Fig. 9. This circuit is called the three-phase diametric (six-phase star). Twelve and occasionally eighteen anodes are used for rectification. The ratio of voltage conversion, that is, volts of direct current to volts of alternating current effective for a few multiphase circuits, is given in Table 1.

TABLE 1

ANODES	3	6	12	18
$\frac{E_{dc}}{E_{rms}}$	1.17	1.35	1.40	1.41

The voltage ratios given for the ideal circuit do not hold for the actual circuit because of the voltage regulation of the transformer and the cathode-anode drop in the rectifying unit. Hence the load voltage reg-

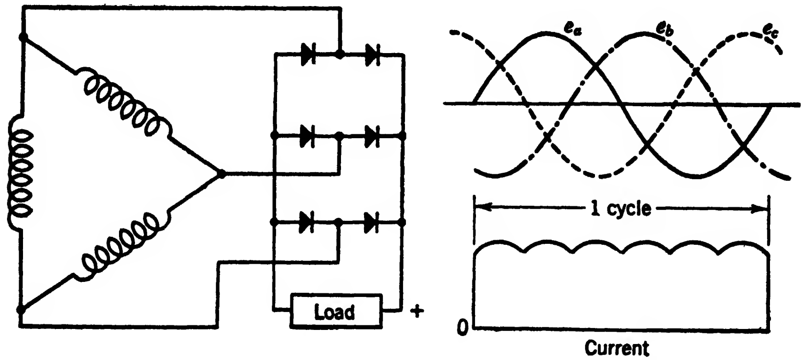


FIG. 8. Full-wave bridge, three-phase rectifier with voltage and current relationship.

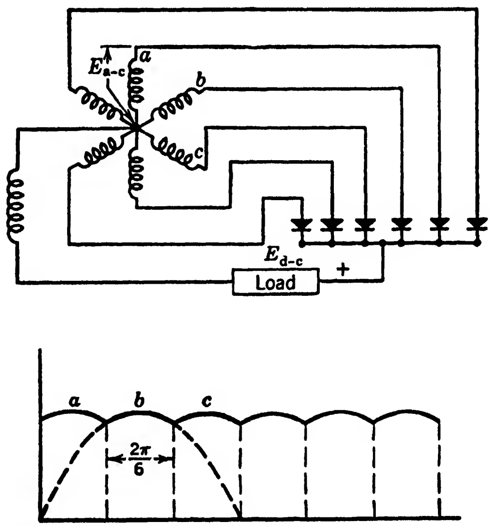


FIG. 9. Schematic six-phase rectifier circuit and wave form.

ulation of the multiphase circuit may be too poor to meet the requirements of an application. Control of the output voltage can be effected in two ways. One method is to control the transformer input voltage by a regulator or by using taps on the primary side of the input transformer. These methods are not well adapted where large amounts

of power are involved. The second method is to control the time of firing of the anodes by grid action. A delay in the firing of each anode will change the wave form of the rectified current and voltage and reduce the magnitude of the average value. The application of this method is suggested in Fig. 10. This method of voltage control increases the ripple and the ripple factor.

In the ideal multiphase rectifier circuit the anode at the highest positive potential carries the entire d-c load. As the potential of the conducting anode falls to that of a succeeding anode, the former drops the load and the latter picks it up so that only one anode is conducting at a time. In the actual rectifier circuit this sudden change of load requiring an infinite rate of change of current for the anodes participating in the action does not take place. The transformer secondary winding that supplies each anode has some leakage reactance which opposes the sudden change of current. Accordingly, the current in the first conducting anode tapers off and the current in the second anode rises as the two anodes pass each other with regard to positive potential. Thus in practice there are periods of transition when two anodes are conducting at the same time. In special cases, such as twelve- or eighteen-anode circuits, there may be three or more anodes conducting simultaneously. A twelve-phase ignitron rectifying unit is shown in Fig. 23.

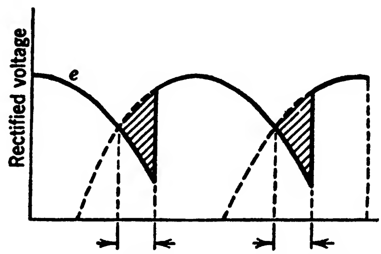


FIG. 10. Voltage control through delay in firing time.

In multiphase rectifier circuits, as in single-phase circuits, the part-time rectification of current from each phase produces distorted (non-sine-wave) currents in the transformer secondaries and may produce residual magnetomotive forces and fluxes in the cores of the transformer feeding the rectifier. These things introduce problems in the design and selection of transformer connections to give optimum results and efficiency. An analysis of these problems lies in the realm of the study of power transformers. However, it should be noted that simple multiple-phase circuits such as shown in Figs. 7 and 9 are seldom employed. The high peak-to-average demand on the rectifying elements together with poor transformer utilization, efficiency, and regulation dictate the use of multiple-conduction circuits using interphase and other special transformer connections like those illustrated

in Fig. 11. Thus a rectifier may sometimes be referred to as a 36- or 72-phase rectifier (common in electrochemical installations), though it is generally made up of a number of three-phase wye groupings properly phased and interconnected with interphase transformers.

The power losses in multiphase rectifiers consist of (1) those in the

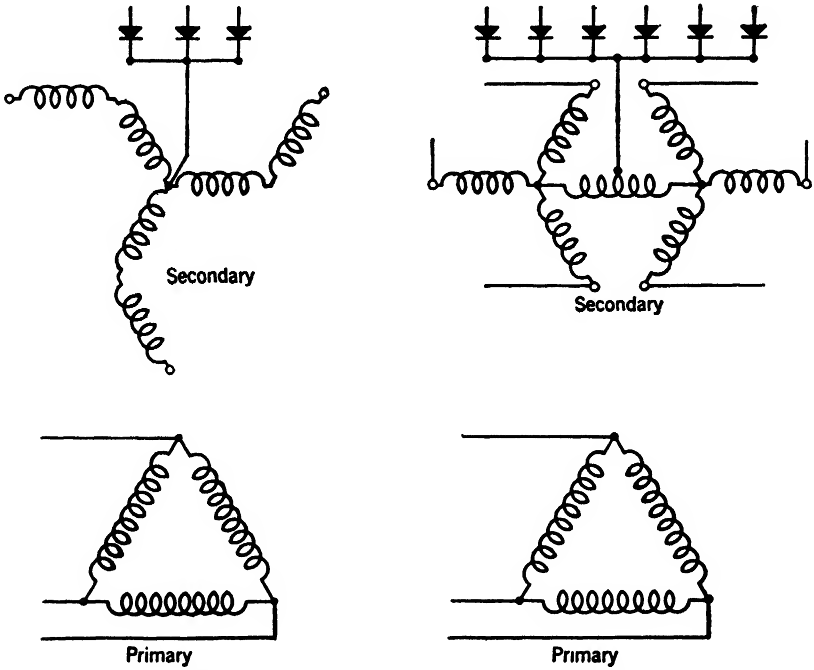


FIG. 11. Delta-zigzag and interphase transformer connections for three- and six-phase rectification.

input transformer, (2) those in the rectifying unit, plus (3) those in auxiliary equipment. Transformer losses under load are relatively low and the full-load efficiency of transformers is very high. In the rectifying units the losses consist of power for producing thermionic emission (if thermionic cathode is employed) plus the loss in the cathode-anode circuit. The cathode-anode fall of potential is of the order of 10 to 20 volts and if the output voltage is 600 volts or higher, the percentage of power lost across the rectifying unit is very low. Thus for 600 volts and an arc drop of 15 volts the power loss in the arc is 2.5 per cent, and for a load voltage of 3000 volts with 15 volts arc drop the power loss is only 0.5 per cent. The over-all efficiency of

complete rectifying units of moderate and large size may well be of the order of 95 to 97 per cent.

Summary on Rectifiers. A summary covering some of the important factors and applications of rectifying units is given in Table 2.

TABLE 2. SUMMARY ON RECTIFIERS

Device	Cathode Content	Load Voltage Range	Load Current Range	Inverse Voltage Peak	Applications
Kenotron (high-vacuum diode)	Hot (vacuum)	0–100,000	Few amperes	100,000	High voltage-low current plate supply for X-ray tubes, radio transformers, Precipitron, etc.
Gaseous rectifier (Tungar)	Hot (gas)	0–220	0–15	300	Charging storage batteries.
Phanotron	Hot (vapor)	0–2,500	0–30	5,000	Charging storage batteries. Generator and motor fields.
Thyratron	Hot (gas or vapor)	0–100,000	0–100	100,000	Control equipment. Lighting, timing, relays, heating, etc.
Grid-glow tube	Cold (gas)	0–300	0–0.0001	600	Control for sensitive relays.
Ignitron	Mercury pool	0–3,000	0–∞	3,000	Power rectification, resistance welding, inversion.
Metal-tank rectifier	Mercury pool	0–3,000	0–∞	3,000	Power rectification, inversion.
Excitron	Mercury pool	0–3,000	0–∞	3,000	Power rectification.
Blocking-layer rectifiers	Cold metal plate	7–15 volts per cell	0–1,000	5–50 volts per cell	Battery charging, electroplating, motion-picture arcs, cathodic protection.

Smoothing Filters. A smoothing filter is a circuit element or a network designed to reduce the ripples of current and voltage in the output of a rectifier. Some applications of direct current can utilize the pulsating output of a rectifier without modification but in the majority of cases a smoothing out of the ripples is necessary. The basic elements for smoothing action are condensers and inductance or choke coils. The former act through their ability to store energy in electric fields and the latter through a similar storage of energy in magnetic fields.

Condensers for filter service are made in several types and ratings. Electrolytic condensers have a thin dielectric film of oxide on aluminum foil. They combine a large capacity with a small size and are available for potentials up to approximately 800 volts. For higher voltages condensers are made with a dielectric of thin paper impregnated with oil. For very high voltages mica is used as a dielectric.

Chokes or inductance units for filters consist of coils wound on iron cores having a short air gap in the core to prevent magnetic saturation at normal load current values. If the iron core approaches magnetic saturation its permeability decreases, the inductance of the unit thus being decreased. The change of inductance with the magnitude

of the current is sometimes useful in filter application. Thus a choke with an air gap in its core may present a high inductance at low values of current and a much lower inductance as magnetic saturation is approached. Such an inductance is termed a *swinging choke.*

The simplest filter circuit for a half-wave, single-phase rectifier consists of a condenser placed in parallel with the load on the output side

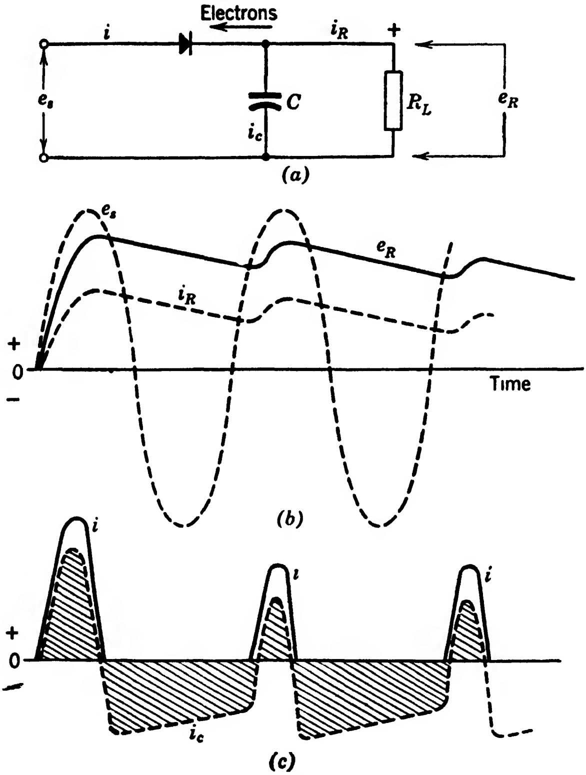


FIG. 12. Half-wave, single-phase rectifier with a simple condenser filter.

of the rectifier as shown in Fig. 12*a*. Assuming (1) an impressed voltage having a sine wave, (2) a rectifying unit (low resistance in forward direction), and (3) a high resistance load R_L, the voltage and current values in the circuit will vary as depicted in parts *b* and *c* of Fig. 12. First, the impressed voltage follows dotted line e_s which will charge the condenser C as it rises to the first peak. Then, as the impressed voltage falls below the voltage across the condenser, the rectifier will cut off and the energy stored in the condenser will discharge through the resistance R_L in accordance with the transient equation:

$$i = -\frac{E^{-t/RC}}{R}$$

The voltage across the condenser and load will decrease, as shown by the full line e_R, until the next positive loop of impressed voltage rises above e_R and then the condenser charges again. The current i_R delivered to the load follows the trend of the voltage across the load.

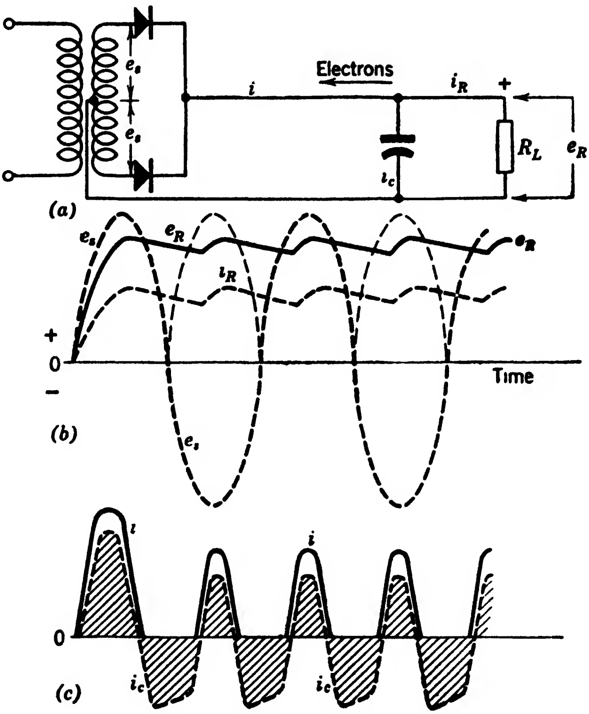


FIG. 13. Full-wave, single-phase rectifier with a simple condenser filter.

The current i flowing through the rectifying unit rises in pulses as shown by the full line in part c, Fig. 12, and the charge and discharge current for the condenser follows the trend of the dotted line i_c. It should be noted that the pulses of current through the rectifier have a high transient peak and will damage the cathode of a gaseous or vapor-rectifying tube. Obviously, the magnitude of the load current i_R and the magnitude of the voltage and current ripples will be determined by the values of C and R_L. In the actual rectifier the rectifying unit has some resistance which produces a voltage drop and reduces the voltage across the load. As R_L approaches ∞, e_R will approach E_{max}.

A simple condenser filter connected in a full-wave rectifier circuit as shown in part *a* of Fig. 13 will produce voltage and current waves as illustrated in part *b*. In comparison with the half-wave rectifier of the preceding paragraph, this circuit has output ripples of double the frequency but of lower ripple magnitude and gives a smoother output. The pulses of input current i and the condenser current i_c are indicated in part *c*.

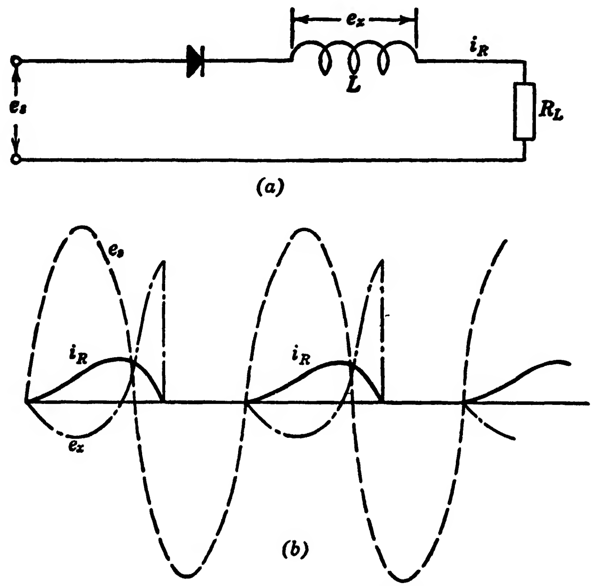


FIG. 14. Half-wave, single-phase rectifier with a choke filter.

A simple choke filter connected in the output of a half-wave single-phase rectifier as shown in part *a*, Fig. 14, will produce the current and voltage waves given in part *b*. The inductance of the choke will retard the rise of current through the rectifier and, after the impressed voltage reaches its positive peak, the energy stored in the magnetic field will continue the d-c pulse after the applied voltage becomes negative. Thus the conduction period will continue for more than one half-cycle, though the rectified current will be extinguished and appear as pulses. This peculiar action is explained by the induced voltage across the choke as illustrated by the curve e_x. The voltage across the load will have the same pulse wave shape as the current. In a full-wave rectifier without filter the current falls to zero at the end of each half-cycle. Here a simple choke filter will serve to reduce the peaks and prevent the current from falling to zero, as shown in Fig. 15. Likewise,

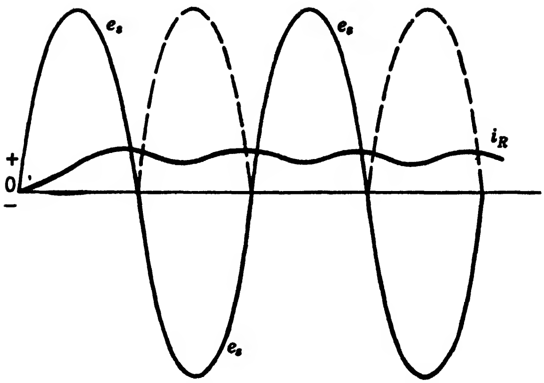


FIG. 15. Full-wave, single-phase rectifier with a choke filter.

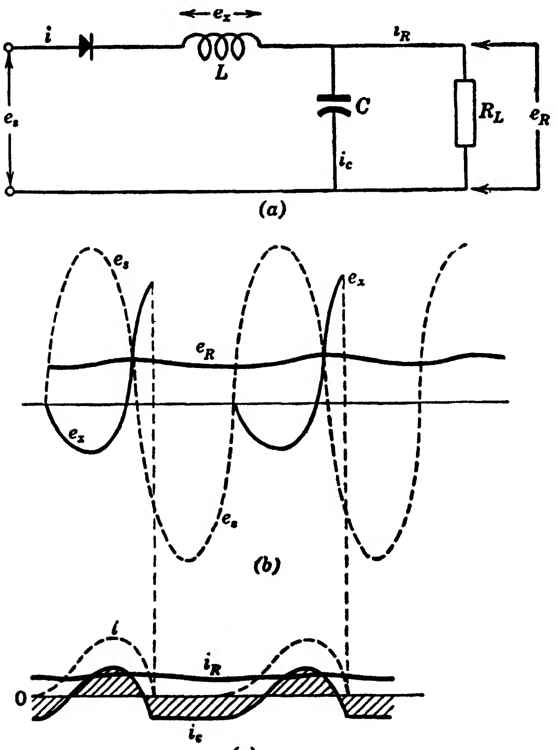


FIG. 16. Voltage and current relations in an L-type filter.

for multiphase rectification where the number of ripples is increased and the ripple factor reduced, the simple inductance filter will be very effective in smoothing out the ripples. The voltage wave form across a resistance load will follow the trend of the current wave.

Filter circuits using combinations of the condenser and chokes will give more effective smoothing action. One simple circuit combination is shown in Fig. 16. Here the series choke will retard the rise and fall of current input to the shunt condenser and load. This action will reduce the magnitude of voltage across the condenser as well as reduce the change in voltage applied to the condenser and load. The condenser, in turn, will absorb energy on the higher voltage inputs and release this energy when the input voltage is low. Both actions tend to maintain a constant voltage across the load and give a good filtering action./ The voltage and current relations in the choke input L-type of filter are shown in parts *b* and *c* of Fig. 16. A second form of L filter uses a shunt condenser on the input side as in Fig. 17. In this circuit the condenser charges rapidly with the rise of rectifier output and to a much higher voltage than in the preceding case. The series choke, in turn, retards the change of current flow to the load and tends to hold constant the voltage at the load. Obviously, the voltage at the load is higher than in the preceding choke input circuit.

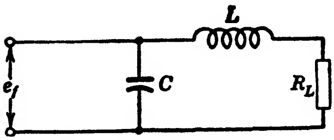


FIG. 17. L-type filter with a condenser input.

The preceding simple filter sections consisting of condensers and inductances in "lumps" are frequently called *brute-force filters.* These filters provide satisfactory filter action for many applications at minimum cost. Better filtering action will be obtained if the capacitance and inductance are broken into smaller units and assembled in π and T sections, as shown in Fig. 18. The cost of the complete filter having multiple units is greater because the cost per unit of capacity or inductance is increased as the capacity of the unit decreases. Hence economy as well as the excellence of filtering action must determine the selection of the filter

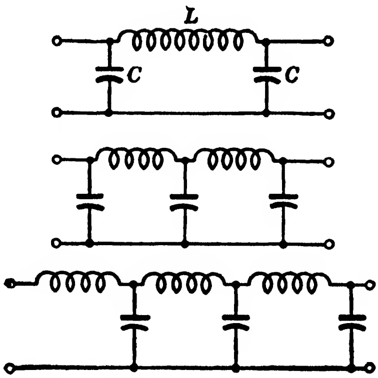


FIG. 18. Combinations of π and T sections to give superior filtering action.

design. In applications where the load current is of low magnitude and where cost or weight of apparatus are critical, resistors are sometimes substituted for chokes in smoothing filter circuits.

Comparison of Filter Circuits. The simple condenser and all other condenser input filter circuits cannot be employed with hot-cathode gaseous and vapor rectifier tubes. This follows because (1) the impedance of a condenser to its charging current is very low, and (2) gaseous and vapor rectifier tubes will not withstand cathode-anode potential drops above approximately 28 volts during the conduction period without damage to the cathode arising from positive ion bombardment. Vacuum-tube and blocking-layer rectifiers can be used satisfactorily with condenser input filters.

Economy and performance determines the selection of the proper filter circuit. The simple parallel condenser is low in cost for half-wave rectification. The condenser input L filter gives a higher load to input-voltage ratio but is likely to give poorer load-voltage regulation. Since the cost of condensers rises rapidly with voltage rating, the choke input may prove more economical for higher load voltages. The choke input L type gives a lower load to input-voltage ratio but has better inherent voltage regulation. The simple series choke filter is usually satisfactory for filtering with multiphase rectifier circuits.

Precautions on Peak Inverse Voltage. In designing rectifier circuits and the accompanying filters, the engineer must give due consideration to the peak inverse voltage to which the rectifying units will be subjected. As pointed out on page 263, the presence of a condenser filter on a half-wave rectifier may subject the rectifying unit to an inverse peak approaching two times the a-c maximum voltage. Transient voltage surges on supply lines or transients due to transformer switching may increase the peak inverse voltages to which rectifiers are subjected. The peak inverse voltage that a tube will withstand (especially gaseous tubes) varies with the operating temperature, and a suitable safety factor should be used for variations of this type.

Analysis and Design of Filters. The current that flows in the output of rectifiers and also in the output of smoothing filters is unidirectional and may be analyzed into a direct-current component and one or more alternating or ripple components. In accordance with fundamental a-c circuit theory, such a current (or voltage) may be expressed by an equation known as a Fourier series. Thus

$$y = f(\theta) = A_0 + A_1 \sin\theta + B_1 \cos\theta + A_2 \sin 2\theta + B_2 \cos 2\theta + A_3 \sin 3\theta + B_3 \cos 3\theta \cdots + A_n \sin n\theta + B_n \cos n\theta \quad (9)$$

If equation 9 is applied to the output of a rectifier the first term A_0 represents the d-c component of current, A_1 and B_1 the component values of the fundamental in the wave applied to the rectifier, A_2 and B_2 components for the second harmonic, and a similar consideration for the other higher harmonics. Usually the value of the constants in equation 9 is unknown but a photographic record of the wave from an oscilloscope is available. For this situation the method of wave analysis described in textbooks on alternating-current circuits may be applied.*

For many rectifier circuits many of the terms in equation 9 are zero or of such small value that they can be neglected. For example, in the idealized half-wave, single-phase rectifier circuit of Fig. 2 the expression reduces to

$$e = E_m\left(\frac{1}{\pi} + \frac{1}{2}\sin\omega t - \frac{2}{3\pi}\cos 2\omega t - \frac{2}{15\pi}\cos 4\omega t\cdots\right) \tag{10}$$

Here the first term (E_{dc}) and the second term, the fundamental a-c component, are the important terms as far as magnitude of voltage or current is concerned. For the idealized full-wave, single-phase rectifier circuit of Fig. 3, equation 9 becomes

$$e = 2\frac{E_m}{\pi}\left(1 - \frac{2}{3}\cos 2\omega t - \frac{2}{15}\cos 4\omega t\cdots\right) \tag{11}$$

For this case the first two terms (I_{dc} and the second harmonic) are primarily responsible for the determination of the magnitude of the instantaneous voltage. More direct and simple determination of the a-c and d-c components was given earlier in the chapter.

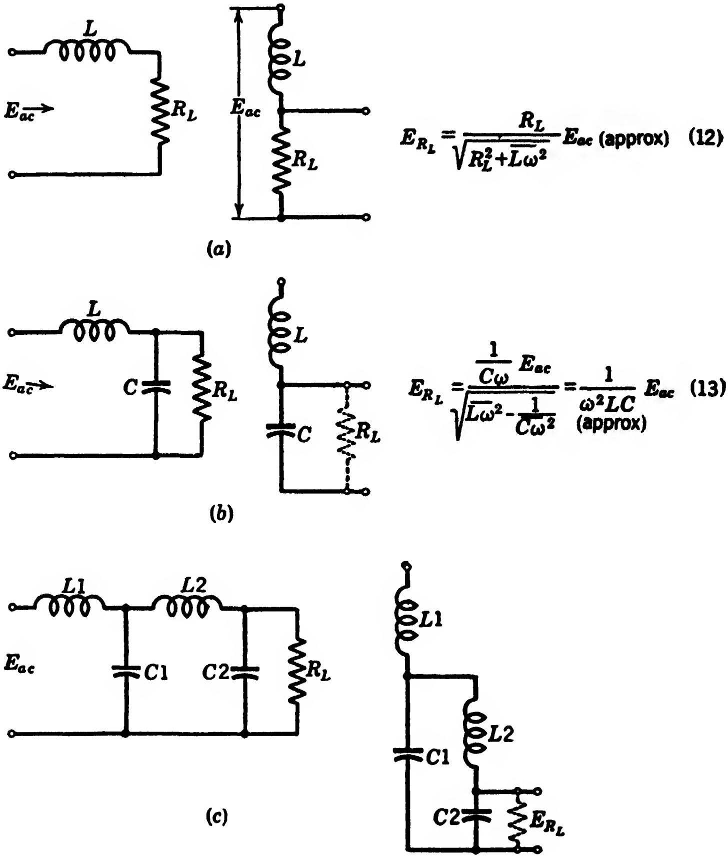
The design of a smoothing filter for a particular application requires the selection of the most economical type of filter and the selection of suitable ratings for the filter elements. The direct objective of the design is to reduce the ripple of the filter output to the magnitude necessary for the application. Ripple is usually expressed in per cent and may be defined as ripple factor times 100. Thus,

$$\text{Per cent ripple} = \frac{\text{a-c ripple voltage (rms)}}{\text{d-c output voltage}} \times 100$$

The a-c ripple voltage in the output of a filter may be determined from the a-c component in the rectifier output by the application of simple

* See Chapter VI, Kerchner and Corcoran, *Alternating-Current Circuits*, second edition, John Wiley and Sons, Inc.

a-c circuit theory. On the principle of the potential divider, this is carried out in Fig. 19. Equations 12, 13, and 14 developed in this



$$E_{R_L} = \frac{R_L}{\sqrt{R_L^2 + L\omega^2}} E_{ac} \text{ (approx)} \quad (12)$$

$$E_{R_L} = \frac{\frac{1}{C\omega} E_{ac}}{\sqrt{L\omega^2 - \frac{1}{C\omega^2}}} = \frac{1}{\omega^2 LC} E_{ac} \text{ (approx)} \quad (13)$$

$$E_{c1} = \frac{1}{\omega^2 L1C1} E_{ac} \qquad E_{c2} = \frac{1}{\omega^2 L_2C_2} E_{c1}$$

$$\text{Combining} \quad E_{R_L} = \frac{1}{\omega^4 L1L2C1C2} E_{ac} \text{ (approx)} \quad (14)$$

FIG. 19. Circuits for approximate calculation of ripple voltage at output of filters. (Equations apply to lowest ripple frequency, not supply frequency.)

figure neglect the resistance of the chokes and assume that the load resistance R_L is large compared with the reactance of the capacitors.

The following approximate equations for the calculation of per cent ripple are taken from the 1945 edition of the *Radio Amateur's Handbook*, page 178.

Single-section filter (Fig. 19*b*) full-wave rectifier $\Big\}$ $\% \text{ ripple} = \dfrac{100}{LC}$ (15) *

Two-section filter (Fig. 19*c*) full-wave rectifier $\Big\}$ $\% \text{ ripple} = \dfrac{650}{L1L2(C1 + C2)^2}$ (16) *

These equations can be used for other ripple frequencies by multiplying each inductance and capacity value in the filter by $120/F$, where F is the actual ripple frequency.

The function of the filter input choke in the three filter circuits of Fig. 19 is to limit the peak output current from the rectifier and to raise the ratio of the average to peak current (smoothing effect). To perform this function the impedance of this choke must be high. The necessary magnitude for this input choke can be calculated approximately for the full-wave, single-phase circuits of Fig. 19, as follows. First, it is obvious that for good filter action the instantaneous current in the input to the filter should never fall to zero. This current is composed of a d-c component and several a-c components, the maximum of which is the fundamental component of equation 11. The d-c component of the current through the input choke is E_{dc}/R_t, where R_t is the sum of the load R_L and the choke resistances. The peak a-c fundamental component of current is $E_{F1}/L\omega$, if $L\omega \gg R_t$ and E_{F1} is the amplitude of fundamental ripple voltage component. Hence for good filter action

$$\frac{E_{dc}}{R_t} \geq \frac{E_{F1}}{L\omega}$$

and

$$\frac{L\omega}{R_t} = \frac{E_{F1}}{E_{dc}}$$

$$L = \frac{R_t}{2\pi f}\frac{E_{F1}}{E_{dc}} = \frac{R_t \cdot \frac{2}{3}}{2\pi \cdot 120}$$

$$= \frac{R_t}{1130} \tag{17}$$

where the ripple frequency is 2×60 and the ratio of E_{F1}/E_{dc} is $\frac{2}{3}$ from the second and first term of equation 11. Equation 17 gives the critical value for L in the input choke. Experiments have shown that a somewhat higher value of L is desirable because of neglected higher

* L is in henries and C in microfarads.

harmonics so that the usual and more conservative formula given in handbooks is

$$L_{(\text{henries})} = \frac{R_L}{1000} \tag{18}$$

There are no simple formulas for calculating the ripple voltage of filters having a condenser input (Figs. 12 and 13), but this voltage will decrease as the magnitude of the capacitor and inductance is increased. Experimental data using the oscilloscope offer a practical solution. Filters for radio power supplies use capacitors of the order of 4 to 8 μf and chokes of the order of 10 to 30 henries for single-phase, full-wave rectifiers having a ripple frequency of 120 cycles. For ripple frequencies other than 120 the inductance and the capacities should be multiplied by the ratio $120/F$, where F is the ripple frequency.

It is considered good practice to connect a high resistance across the output of a filter to bleed off the condenser charge when the power supply is not in service. This resistance known as a *bleeder resistance* should be selected to draw 10 per cent or less of the rated output of the power supply.

The output voltage of a filter falls with an increase of load. At no load the output voltage approaches the maximum of the impressed a-c wave. Normally the bleeder resistance will draw sufficient current to hold the output voltage below the a-c maximum. As the load on the power supply is increased the output voltage will fall due to the voltage drops in (1) the rectifying unit, (2) the resistances in any series chokes, and (3) the resistance of input transformer windings. The operation of the choke input filter is improved by a "swinging choke" for the input unit. This choke provides a high value of inductance when the only load consists of the bleeder resistance and then gives a lower inductance as the load rises (R_L decreases) to fulfill the requirements of equation 18. Thus the bleeder current is the critical current which determines the rating of the input choke. In making calculation of the voltage regulation of a power supply the current taken by any bleeder resistance must be added to that passing through the load resistance.

Inversion Circuits. An inverter is a device for converting direct current to alternating current. The oscillator discussed in Chapter IX is such a device and is designed for high frequencies and generally for small amounts of power. For low frequencies and larger amounts of

power, electronic devices such as the thyratron, the ignitron, and mercury-arc rectifiers are used in inverter circuits. The function of the tubes is to commutate or perform a switching operation. It is essential that the electronic device have grid control and that units of inductance or capacity or both be employed to produce inductive and capacitive storage capacity. Many combinations of these elements may be used for producing inversion.

One simple inverter circuit in which the device determines its own frequency is given in Fig. 20. In this circuit the resistance R_1 is relatively high and R_2 is relatively low. Before the d-c switch S is closed,

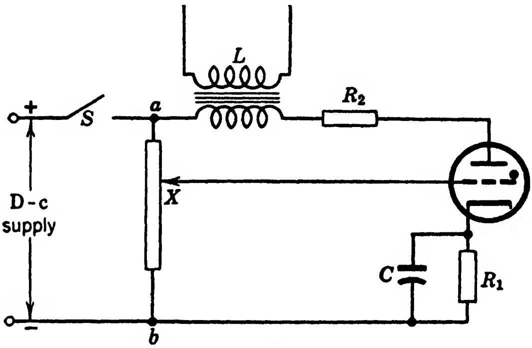


FIG. 20. Simple circuit using a thyratron for inversion.

the anode, grid, and cathode of the thyratron are at the same potential and the condenser C is discharged. When S is closed the grid is made positive with respect to the cathode, and hence if the d-c voltage is higher than the ionization potential, the thyratron will "fire" and current will flow through L, R_2, C, and R_1 in parallel. Since R_1 is large, the condenser will charge quickly and the potential of the cathode will rise. The rise of potential of the cathode will tend to lower the potential across the cathode-anode circuit below the ionization potential. At the same time the inertia effect of the inductance L tends to maintain the current and will produce an induced voltage to maintain the current. As a result the condenser C will continue to charge and the voltage across it will rise above the d-c impressed voltage. When the energy stored in the core of L is expended, the voltage across the thyratron falls, leaving the potential of the cathode higher (owing to condenser) than the anode and higher than the grid. Hence the current through the thyratron ceases and deionization takes place. In the meantime, the condenser discharges through the resistance R_1. As the condenser discharges and lowers its potential and

that of the cathode to the point where anode-cathode potential reaches the ionization potential, and when the grid voltage reaches the point where it will fire the thyratron, the thyratron conducts current again, and the cycle is repeated. Obviously, the time of firing and the resulting frequency of the oscillating cycle can be varied by the position of X on the potentiometer ab which controls the grid bias. The frequency can be controlled also by variation of R_1 or C.

A transformer placed in the circuit of the thyratron would deliver alternating current to a suitable load. The wave form of this type of inverter is not sinusoidal.

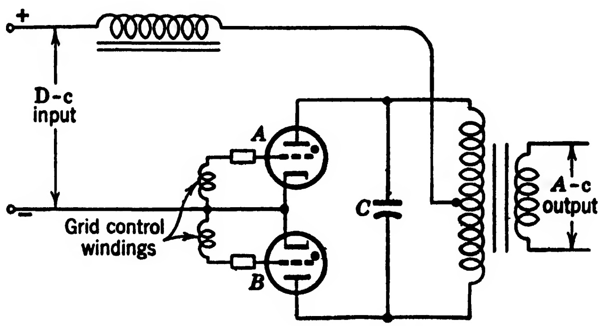


FIG. 21. Single-phase inverter using two thyratrons.

In most power inverters the frequency is determined by a separate source. A simple schematic circuit for a single-phase inverter of this type is given in Fig. 21. Two thyratrons, A and B, with grids controlled by a separate voltage source are employed in a balanced circuit arrangement wherein both anodes are connected to the positive side of the d-c supply. Assume that the grid supply at the instant of starting is positive for A and negative for B. Tube A fires and conducts electrons from the negative d-c line through the upper half of the transformer primary to the positive d-c terminal. The rise of current through the transformer winding creates an IX drop so that the potential of the anode of A falls to a value above zero equal to the cathode-anode drop in the tube. Under this condition the capacitor becomes charged, with the top plate negative and the lower plate positive. When the grid control potential reverses and makes grid of tube A negative and B positive, quick action follows. The negative grid on A has no direct effect upon tube conduction, but the positive grid on B causes it to fire and the resulting surge of current through the lower half of the transformer winding lowers the potential of anode B.

Simultaneously, electrons are supplied to the lower positive plate of the condenser C. This permits the unbound electrons on the top plate of C to flow off and lower the potential of the anode of A, thus permitting the now negative grid of A to regain control. At the end of the half-cycle the reverse process will take place with tube A conducting and tube B cut off. Thus the two tubes A and B serve as a reversing switch

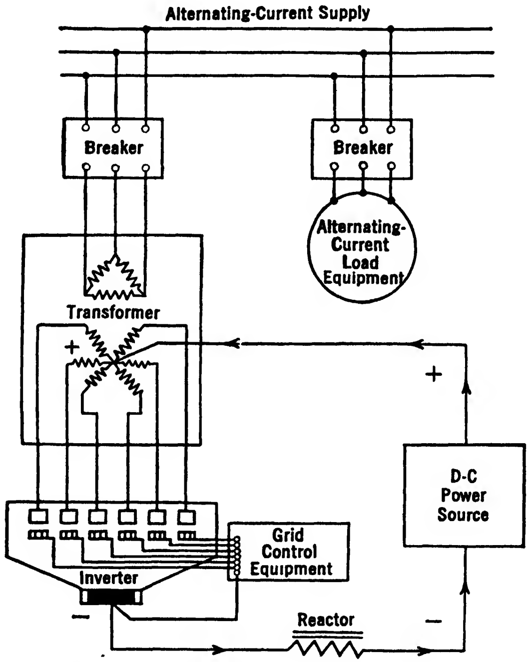


FIG. 22. Circuit using a mercury-arc rectifier for power inversion. (Courtesy Allis-Chalmers Manufacturing Company.)

to cause alternate pulses of current to flow through the primary of the transformer. The resulting flux reversals in the iron core of the transformer gives an a-c output in the secondary having the same frequency as the applied grid potential.

Thyratrons have been used successfully for inversion by the General Electric Company on an experimental d-c transmission line between Mechanicville and Schenectady, New York. The inverter installation consisted of two six-phase units operating in a twelve-phase relationship. The inversion unit had a rating of 5250 kilowatts and took constant current at 28,000 volts at full load from a d-c transmission line. This inverter installation was made in 1936 and served to

convert many millions of kilowatthours of electrical energy until operation was discontinued in 1947.

The multiphase mercury-arc rectifier also can be used as an inverter. A schematic diagram for this application is given in Fig. 22. This diagram is similar to that for rectification of alternating current but with three fundamental differences. First, the polarity at the d-c source is reversed though the direction of current flow through the recti-

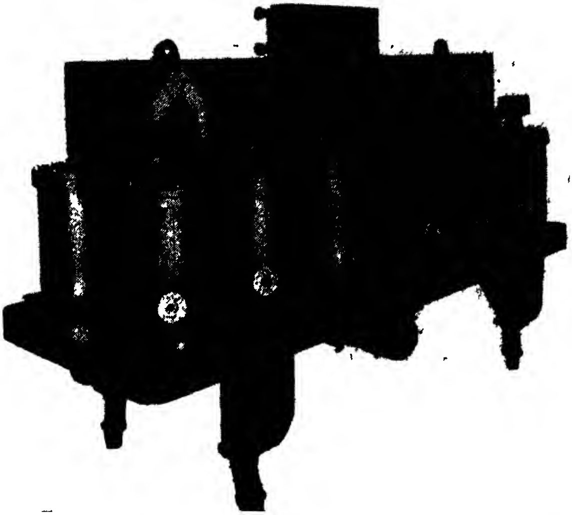


FIG. 23. Twelve-phase ignitron rectifier unit. (Courtesy Westinghouse Electric Corporation.)

fier is the same. Second, a reactor has been placed in the cathode lead. Third, the potential across the grid is controlled by a separate a-c source. The key to the inverter action lies in the application of and control of grid potentials. A positive potential on a grid will attract electrons to the grid and cause them to pass through to the anode. Hence if a-c potentials are applied to the multielectrode grids in the proper sequence, the resulting currents will pass through the secondary windings of the transformers connected to the anodes and produce a-c power in the primaries. The reactor in the cathode lead smooths out the power flow from the d-c source and tends to reduce short-circuit current if an occasional short circuit should occur.

Ignitrons have been used for inversion and are destined to have an extensive application in this field. The circuit employed will follow that for the multiphase mercury-arc rectifier except that special firing

circuits will be needed for the ignitrons. A twelve-phase ignitron rectifier which can be used for inversion is illustrated in Fig. 23.

Possible applications of the mercury-arc rectifier as an inverter are for d-c transmission of power and conversion of d-c regenerated power on electrified railroads in mountainous districts.

Problems

1. In an ideal half-wave rectifier circuit the impressed rms voltage is 65 volts. What is the d-c average voltage? If the peak value of current on the input side is 5 amperes, what is the I_{dc}? I_{rms}?

2. In an ideal full-wave rectifier the rms values on the input are 90 volts and 6 amperes. Calculate the readings on a d-c voltmeter and ammeter in the load circuit.

3. The unfiltered current from a half-wave rectifier (sine wave applied) is passed through a d-c ammeter and an a-c ammeter. When the d-c meter reads 12 amperes, what should be the indication on the a-c meter? Recalculate the problem for full-wave rectification.

4. Recalculate Problem 3 when a current transformer (1/1 ratio) is inserted between the a-c ammeter and the line.

5. A certain phase-shift voltage control on a three-phase rectifier delays the firing time to the peak of the impressed wave. What will be the resulting ripple factor? Refer to Fig. 7 and neglect the effect of inductance in transformer and the arc drop.

6. The vacuum diode of Fig. 15, Chapter III, supplies a rectified current of 100 milliamperes to a resistance load of 2000 ohms. Calculate the ratio of the reduction of the voltage (equation 8).

7. A Tungar rectifier tube described on page 265, is used for half-wave rectification as shown in Fig. 2 to charge a 6-volt storage battery for automobiles. If the battery takes 6 amperes at 7.5 volts, calculate the impressed a-c volts when the tube fires. If the transformer has an efficiency of 80 per cent for this load, calculate the over-all efficiency of rectification, taking care to include both cathode and arc losses. (The student should note that a storage battery is not a resistance load and that a new analysis and approach to the problem is required.)

8. Substitute a copper oxide rectifier for the Tungar tube and recalculate Problem 7, assuming a peak inverse voltage of 12 volts per plate with a 1-volt drop in the forward direction.

9. Two phanotrons of Fig. 7, Chapter XII, are connected in the full-wave rectifier circuit to charge a 110-volt storage battery (125 volts when charging) at 40 amperes. Determine the approximate applied a-c volts on the secondary. If the transformer has an efficiency of 90 per cent, what is the over-all efficiency of rectification?

10. In the simple filter circuit of Fig. 12, assume that $C = 1$ microfarad, $R_L = 100{,}000$ ohms, and e_R falls for $\frac{1}{80}$ second. Determine from Fig. 9*b*, Chapter V, the per cent of voltage drop from the maximum. Redetermine for $R_L = 10{,}000$ ohms.

11. A full-wave rectifier supplies power to a 4500-ohm resistor. The transformer voltage is 300 volts (rms) each side of center tap. Assume that each diode has a constant resistance of 450 ohms when conducting and that the total transformer resistance is 100 ohms. Calculate (*a*) the d-c load voltage; (*b*) the d-c load current; (*c*) the a-c load voltage; (*d*) the per cent ripple.

12. If a filter is added to the power supply of Problem 11, what is the smallest inductance that the input choke may have if the load current is never to fall to zero?

13. A single-section *LC* filter is to be used with the power supply of Problem 11. Suggest suitable values if the ripple is to be 5 per cent and give reasons for your choice.

14. Suggest a suitable filter for the power supply of Problem 11 if the ripple is not to exceed 0.25 per cent.

15. Design a full-wave rectifier delivering 1000 volts d-c at a rated full-load current of 400 ma. Assume that the total internal impedance of the power supply is 200 ohms. The ripple should not exceed 1 per cent. Calculate (*a*) transformer voltage; (*b*) bleeder resistance and power rating; (*c*) inductance limits of input (swinging) choke; (*d*) suggest a suitable filter design; (*e*) indicate the necessary voltage and current ratings for the components.

REFERENCES

LIVINGSTON and LORD, "The Single-Tube Thyratron Inverter," *Electronics*, April 1933.

DURAND, S. R., *A Mercury-Arc Static Inverter*, Allis-Chalmers Mfg. Co., 1939.

M.I.T. ELECTRICAL ENGINEERING STAFF, *Applied Electronics*, John Wiley & Sons, 1943.

Chapter XV

PHOTOELECTRICITY

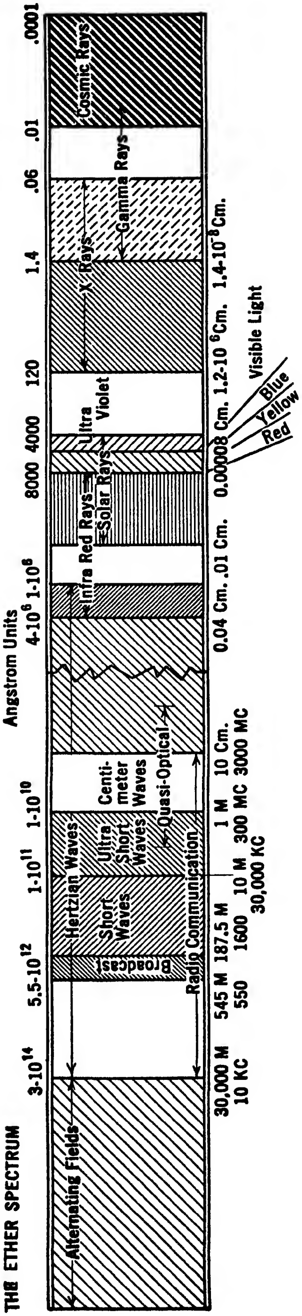
In 1905 Einstein suggested that light consists of closely packed bundles of energy which he termed "light quants." These light quants or darts of light could travel long distances and still maintain the same quantity of energy, enough to emit electrons from a solid. This theory has come to be accepted as the quantum theory of light and of radiant energy. The energy bound up in the light quant is commonly called the "photon." The concept of the scientist today regarding light is a dual theory—light is a form of wave energy and a corpuscle (photon) at the same time.

It is well to remember the reversible energy relationship in our electron theory of today. Thus in a gaseous conduction tube a flying electron may collide with a molecule of gas and give up its kinetic energy in exciting that molecule and cause it to give radiant light energy—in photons—and again a photon of light impinging upon a surface may eject an electron into space.

Spectrum of Radiant Energy. All forms of radiant energy are propagated at the velocity of light which is 3×10^{10} centimeters per second in free space. Radiant energy is assumed to travel in the form of waves and the number of waves per unit of time is called the frequency. Since the rate of propagation is fixed, it follows that

$$\text{Frequency} \times \text{wavelength} = 3 \times 10^{10}$$

This equation gives a simple relation for determining either the frequency or the wavelength when one value is known. Wavelengths may be expressed in meters, centimeters, or angstrom units. The angstrom unit is 10^{-8} centimeter or 10^{-5} micron and is commonly used for radiant energy, particularly in the visible or near-visible range. A chart showing the entire spectrum of radiant energy is given in Fig. 1. The energy in a wave depends on the frequency. Thus cosmic rays are more effective than the gamma rays which, in turn, are stronger than the X rays. The range of radiation visible to the human eye covers but a



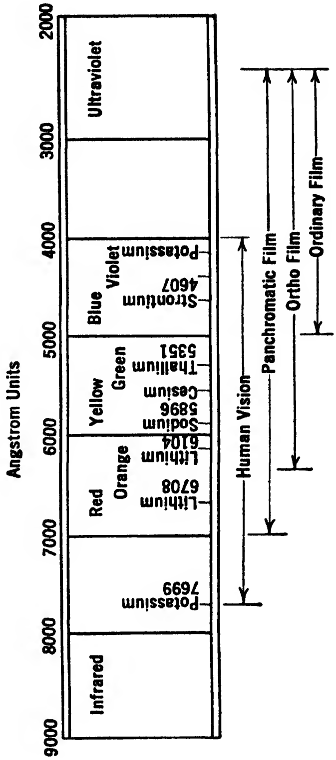


FIG. 1. Spectrum of radiant energy and visible spectrum. (Courtesy *Electronics*.)

very small part of the spectrum (approximately from 4000 to 8000 angstrom units). Below the visible rays are the heat rays which, in turn, lie above a wide band of radio waves. The visible spectrum is shown in the lower right-hand corner of Fig. 1.

Photoelectric Action. The term photoelectric implies the action of light in producing electricity. Since all electrical phenomena involve a displacement of electrons, the term photoelectric means the action of light upon electrons. When photons of light impinge upon solids they have the power of disturbing or releasing the electrons within that solid. The resulting movement of electrons gives rise to various kinds of electric phenomena. Thus electrons may be emitted from a solid and the action is termed *photoemission.* The electrons may be moved across a barrier in the solid, producing a difference in potential, giving rise to a *photovoltaic* action. Lastly, the photons may penetrate the solid sufficiently to release electrons and change the resistance of the substance, giving rise to the term *photoconduction.* Commercial devices have been built to utilize these three different aspects of photoelectric action.

Photoemission. In 1888 William Hallwachs found that, if he charged a zinc plate to a negative potential and then exposed it to ultraviolet light, it gradually lost its charge. However, the plate did not lose its charge when exposed to ultraviolet light after being raised to a positive potential. This phenomenon has been thoroughly investigated since the time of Hallwachs and it has been found that metals and many other substances as well exhibit it to a greater or lesser degree. This phenomenon wherein the electric charge upon a body may be changed by light is known as the Hallwachs effect.

Two famous co-workers, Julius Elster and Hans Geitel, experimented with many metals and observed that the photosensitivity of aluminum, magnesium, and zinc was superior to that of many other metals. From this they reasoned that the more electropositive metals would give better results. This expectation was proved by experiments with the alkali metals, sodium and potassium, and later with amalgams of these metals placed in an evacuated tube. Later experiments by others have shown that cesium, rubidium, and the alkaline earths, particularly strontium and barium, are satisfactory photoelectric emitters.

The theory of photoelectric emission is analogous to that of thermionic emission. Photoelectric emission results from the kinetic energy of the photon of light being imparted to the electrons in the surface of the emitter. In thermionic emission the kinetic energy of agitation due to the high temperature of the cathode gives some electrons at the

surface sufficient velocity to overcome the electron affinity at the surface. Satisfactory photoelectric emitters have a low work function. According to the Bohr theory, their atomic structure consists of a nucleus surrounded by a closely grouped band or completed system of electrons plus one lone outer electron in alkali metals, and two lone electrons in alkaline earths. It is these lone electrons which are subject to removal. The photons of light meeting these lone electrons may add to the normal velocity of the electrons sufficiently so that they overcome the electron affinity or work function and fly away from the emitting substance.

There are two laws of photoelectric emission. The first states: *The number of electrons released per unit of time at a photoelectric surface is directly proportional to the intensity of the incident light.* This law has been tested for a range of intensities varying from zero to full sunlight. Wherever apparent deviations have been discerned they can be explained by errors in measurement or by inherent faults within the cell employed which tended to introduce spurious currents or to hinder the total number of electrons actually released from being collected. This law makes the principle of photoemission very valuable for light measurement and many other applications. The second law states: *The maximum energy of electrons released at a photoelectric surface is independent of the intensity of the incident light but is directly proportional to the frequency of the light.* This law implies that the energy imparted to the electron by electromagnetic radiations is directly proportional to the frequency of these radiations.

The reason for the first law of photoemission is rather obvious but that for the second law is not so clear. The second law was explained in 1905 by Einstein who reasoned that incident radiant energy was transferred to surface electrons in a quantum of magnitude hf wherein a portion of this energy was used for removal of the electrons and the remainder appeared as kinetic energy. In mathematical form Einstein states the relationship as follows:

$$hf = \phi + \tfrac{1}{2}mv^2 \tag{1}$$

where h is Planck's constant, f is the frequency of the incident light, and ϕ is the work function of the emitter in equivalent electron-volts. The term hf is the energy of the impinging light (photons) which overcomes the work function ϕ and gives a velocity v to the emitted electrons. Inspection of equation 1 shows that if hf equals ϕ the electron will have zero velocity and will not be emitted. Thus there is a mini-

mum or *threshold frequency* for the incident radiation below which photoelectric emission will not result. Equation 1 has been verified experimentally. The second law of photoemission follows from equation 1 because h and ϕ are constants for a given emitter, making $\frac{1}{2}mv^2$ proportional to f. Equation 1 holds only for the electrons lying on the surface which will have the maximum velocity of emission. Some light will penetrate the outer layer of atoms and liberate electrons within the emitter, but these electrons are likely to lose much of if not all their energy before leaving the surface.

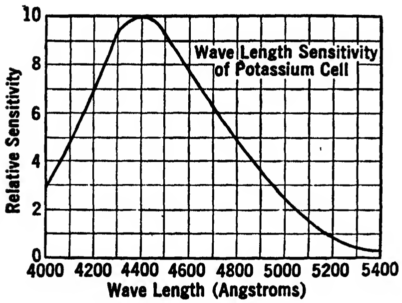


FIG. 2. Photoemissive color sensitivity of potassium. (Courtesy General Electric Company.)

Color sensitivity is the relative response of photoemissive materials to various light radiations. A curve of the sensitivity of potassium is given in Fig. 2. This curve covers the green-blue-violet part of the visible spectrum with the peak occurring in the violet range. Photosensitive materials having a peak in the red range of the spectrum are very desirable for use under the light from incandescent lamps which radiate their maximum energy in the near infrared.

Photosensitive films generally show a much greater emission than a pure metal or substance. In this respect and several others, photoelectric emission parallels thermionic emission. Thus thermionic emission from tungsten is greatly increased by a film of thorium one molecule thick, and emission from a metal covered by barium oxide with a monomolecular layer of barium on the outside is still greater. The surface treatment of thermionic emitters also greatly affects the resulting emission. These various factors control the work function of the emitting surface. In a similar manner the photoelectric emission may be improved by a thin (molecular) layer of cesium on magnesium, whereas a much greater emission may be attained by a thin deposit of cesium upon a sub-base of cesium oxide covering a base of pure silver. Again, the treatment of potassium in hydrogen will change the color sensitivity and will improve the emission of that light-sensitive material.

Vacuum Phototubes. The photoemissive tube consists of a cathode having a photosensitive surface and an anode placed in a glass envelope. These two electrodes are connected in series with a battery

(anode positive) and a load consisting of a resistance as shown in Fig. 3. Light falling upon the cathode causes the emission of electrons which are attracted to the positive anode and which cause a current through the external circuit including the load. In the vacuum phototube the space within the glass envelope is highly evacuated.

Many different types of construction have been used for phototubes. In the early type the cathode frequently consisted of a deposit of light-sensitive material on the inner wall of the spherical tube. A circular window for admitting light was made on one side of the sphere by evaporation using the heat of a Bunsen burner. Anodes usually consist of a straight wire or wire ring placed near the center of the envelope. Sodium, potassium, barium, rubidium, and cesium have been the materials used for the light-sensitive surface. Nearly all commercial tubes produced today are enclosed in cylindrical glass envelopes as shown in Fig. 4. The cathode is a plate bent into the form of a semicylinder. The light-sensitive material is deposited on the cathode after the tube is assembled and evacuated. For the measurement of ultraviolet light either sodium or cadmium may be used. In order to admit the ultraviolet light, quartz or some special glass such as Corex must be used. One scheme for admitting ultraviolet is to use a special thin circular window (1-inch diameter) of Corex glass which projects out from the standard cylindrical glass envelope. The highest sensitivity of phototubes to light from incandescent lamps is obtained through the use of the cesium oxide tube (developed by L. R. Koller) which consists of a monomolecular coating of cesium on a sub-base of cesium oxide upon a base of silver. This type of tube may be produced in a number of ways, one of which is as follows.*

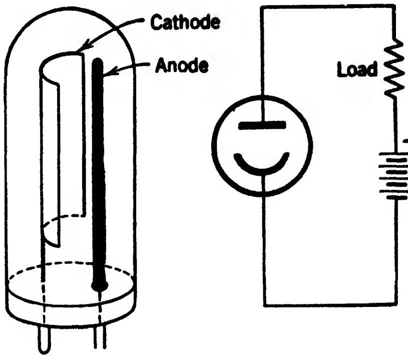


Fig. 3. Construction and circuit of a phototube.

The cathode is made of solid silver or of silver-plated copper in the form of a semicylinder. The anode is a coaxial vertical wire with a disk of nickel welded to its top. The disk holds in a pocket a small pellet of cesium chloride and calcium or of cesium dichromate and

* The following description is taken from *Photocells and Their Application*, by V. K. Zworykin and D. Wilson, John Wiley & Sons, 1934.

silicon. The glass press and its electrodes are sealed into a cylindrical glass bulb. The tube is exhausted and baked in a furnace. After cooling, about 2 millimeters of oxygen are admitted and a glow discharge is passed between the anode and cathode. Oxidation of the silver is evidenced by a progressive series of brilliant interference colors on the cathode. The glow is continued intermittently until the color of the cathode passes through the first bright green. The residual

General Characteristics		
Vacuum and gas diodes	Vacuum	Gas
Cathode surface	S_1	S_1
Cathode size	0 81 x 1.38	0.81 x 1.38 in.
Luminous sensitivity, 0 cycles, μamp/lumen	15	50
Luminous sensitivity, 10,000 cycles, μamp/lumen	15	55
Radiant response μa/μW at 7500 Au	0.0015	0.0060
Response max at	7,500	7,500 A
Response, upper limit	11,000	11,000 A
Response, lower limit	4,000	4,000 A
Gas amplification, max	.. .	7
Capacitance, anode-cathode	2.5	2.5 μμf
Leakage resistance, min	4,000	90 meg
Mounting position	Any	Any
Net weight	2 oz	2 oz
Shipping weight	8 oz	8 oz
Maximum Ratings		
Anode voltage, peak	500	90 volts
Anode current	20	20 μamp
Temperature, ambient air	100°	100° C

FIG. 4. Phototube and characteristics. (Courtesy Westinghouse Electric Corporation.)

gas is then pumped out. By means of a high-frequency field, the pellet disk is heated to a temperature high enough to explode the pellet and liberate cesium vapor which initially condenses on the glass walls of the bulb. The tube is then placed in a furnace and held at a temperature of 200 to 225 degrees C. If the back of the cathode or the stem of the press is painted with a mixture of lead oxide or tin oxide in amyl acetate, all excess cesium will be absorbed by the paint and the cathode will become dark gray in color as the baking proceeds. When the edges of the cathode begin to turn light in color, the tube is completed. The color-sensitivity curve of a cesium oxide vacuum phototube made in this manner is given in curve S_1 of Fig. 5. The sensitivity of this tube is of the order of 10 to 20 microamperes per lumen.

Since the current output of vacuum tubes is so small, any leakage between the cathode and anode lead-in wires may be troublesome. For low light intensities any leakage current through the glass or on

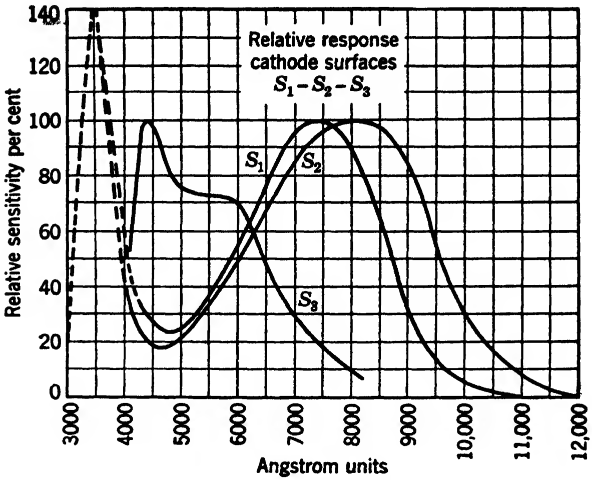


FIG. 5. Relative responses of three different photosensitive surfaces (dashed lines are outside the range of human vision) S_1 is cesium-oxygen-silver, S_3 is rubidium-oxygen-silver, and S_2 is discontinued. (Courtesy Westinghouse Electric Corporation.)

the inside or outside walls of the envelope may mask the real emission current. The magnitude of the leakage current may be reduced in two ways. If the leads come to the tube base, a conducting ring may be sealed in the glass stem around the anode lead and connected in shunt around the load resistance as shown in Fig. 6. A simpler arrangement is to bring the lead of one electrode out at the base and the other out of the top of the tube, making a long path for any possible leakage current.

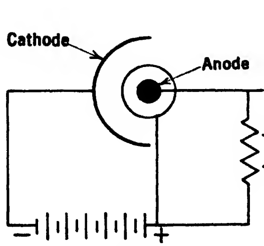


FIG. 6. Circuit for eliminating leakage current in the load of a phototube.

Negative space charge has little influence in the vacuum phototube because the cathode is made relatively large to collect light and the emission is small. Some space charge may exist near the anode but if this electrode is maintained strongly positive, the space charge has no detrimental effect.

The variation of the anode current with anode potential for constant light intensity and a given color value is shown by Fig. 7. This indicates that a low potential of 20 to 40 volts will attract nearly all elec-

trons emitted. Potentials of 90 to 250 volts are used in the circuits of vacuum phototubes. The resulting cathode-anode voltage drops for four values of load resistance may be obtained at the intersections of the load lines with the characteristic curve.* The curve for anode current versus light flux is a straight line following the first law of photoemission. Any deviation from this characteristic is due to fatigue in the light-sensitive material or some variable electric leakage.

The vacuum phototube is suited for applications where the use of a

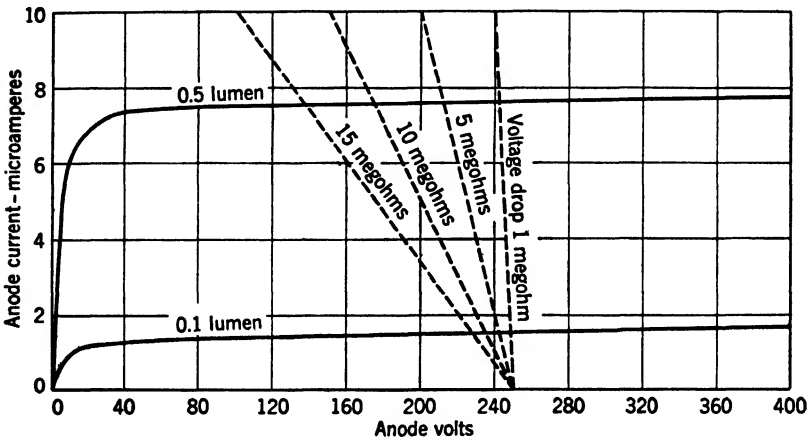


FIG. 7. Anode-current voltage characteristic of a vacuum phototube. (Courtesy Westinghouse Electric Corporation.)

high resistance load is desirable to give maximum circuit sensitivity. It offers a high stability of operation and permanence of calibration. Thus it is used for light-measuring and light-relay devices. This tube is also used where good dynamic response † is needed.

Gaseous Phototube. The admission of a small amount of inert gas into a vacuum phototube will greatly increase the current flow owing to light falling upon the cathode. This amplification of current or increase in sensitivity may be tenfold in value, but it is usually limited to the order of three to seven times the current in a vacuum. This increase of current parallels that found in thermionic tubes containing gas, but the cause of the phenomenon is somewhat different. In the thermionic tube the increase is due to the neutralization of negative

* The load line represents the voltage drop across the load resistor of Fig. 3. Subtracting this drop from the d-c supply voltage gives the voltage remaining across the phototube.

† See page 374 for explanation of dynamic response.

space charge near the cathode, whereas space charge is of little consequence in the phototube. The reasons for the current increase in the gaseous tube are (1) the additional ionic current which results as soon as the ionizing potential is reached; and (2) an increase in emission of electrons from the cathode due to the bombardment by positive ions. The change in anode current with rise in anode potential is given in Fig. 8 for five different amounts of light flux in lumens falling on the cathode. Obviously, a low anode potential attracts all the primary

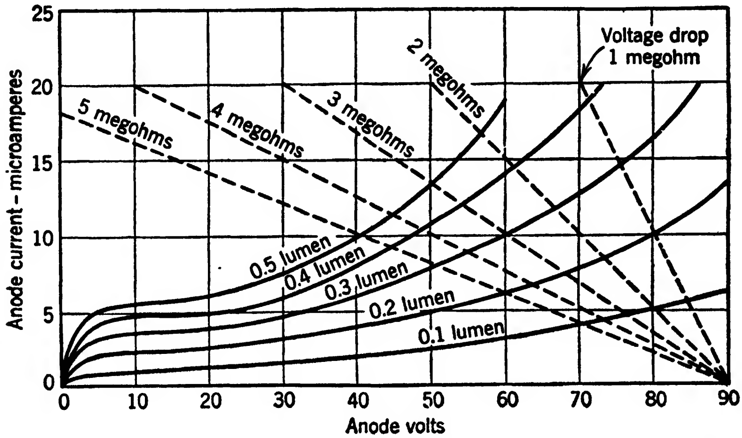


FIG. 8. Anode-current voltage characteristic of a gaseous phototube. (Courtesy Westinghouse Electric Corporation.)

electrons emitted. At 20 volts (approximately) the ionizing potential is reached and beyond that the current increases with voltage, at first slowly, later becoming accumulative in effect following the theory of the Townsend discharge described in the chapter on gaseous conduction. The potential applied to the gaseous phototube circuit is 90 volts and the actual cathode-anode drop is limited by the load resistance. The trend of the curve will depend on the gas pressure and other factors. The operation of the gaseous tube must be kept below the glow-discharge point or the control of its operation by incident light is lost.

The light-flux-current characteristic of the gaseous tube is given in Fig. 9. This curve is not linear but the deviation does not cause any undesirable distortion of music or voice when the tube is used for sound on film.

One disadvantage of the gaseous phototube is its poor dynamic response. Dynamic response corresponds to the high-frequency response

of amplifier tubes and,circuits. It refers to the speed with which the anode current follows changes in the light incident on the cathode. It is important when there is a high frequency in the light changes. The poor dynamic response is due to the current conduction by ions in this cell. At a given instant the space between the cathode and anode contains a large number of electrons and positive ions. The electrons are swept out of the space instantly by the positive potential on the

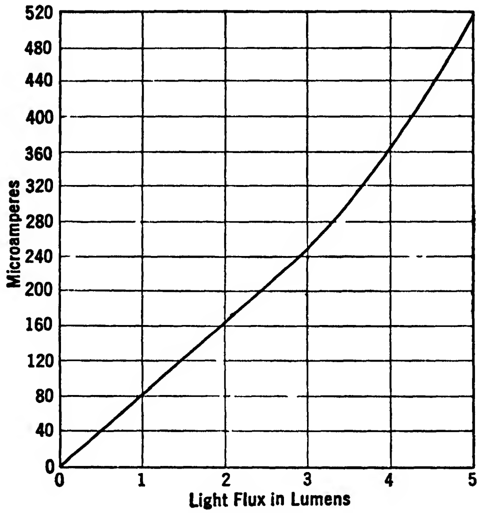


FIG. 9. Light-flux-current characteristic of a gaseous photocell.

anode but the positive ions, being relatively ponderous, move much more slowly toward the cathode. If the light drops to zero at a given instant, the positive ions take a little interval of time to reach the cathode; when they do reach it they may release a few electrons which then dart to the anode. Thus, although the light goes to zero instantly, the anode current lags in making the change. If alternating changes in light flux take place, the anode current will fall behind in making like changes and the lag will be proportional to the frequency of that change. The difference in the dynamic response between the vacuum unit and the gaseous unit is illustrated in the curve of Fig. 10.

The standard commercial gaseous phototubes use a cesium oxide cathode. They are filled with argon at a pressure of about 100 microns. (A micron is a pressure equivalent to a column of mercury one-millionth of a meter high.) They have a sensitivity of about 50 microamperes per lumen though some are built with sensitivity of 60 to 100.

The more common commercial tubes are about 1⅛ inches in diameter and 4⅛ inches high (including base and pins). The cathodes are about 1⅜ inches high with a diameter of 11/16 inch.

Curves of the color sensitivity of three types of cathode surfaces are given in Fig. 5. Surface S_1 covers the spectrum of the incandescent lamp satisfactorily. Surface S_2 is adapted for infrared rays and surface S_3 works best in the blue and ultraviolet region of the spectrum.

Gaseous phototubes are used almost universally for sound reproduction from films. They serve well in many other applications not requiring a high dynamic response. Some of these applications are: (1) for facsimile transmission of pictures, (2) as an "electric eye" in photometric measurements, (3) as an "electric eye" in color analysis and color matching, (4) for calibration of watthour meters, (5) for miscellaneous counting operations, (6) for detection of defects in manufactured articles, (7) for control of sign lighting, and (8) for control of artificial illumination in factories.

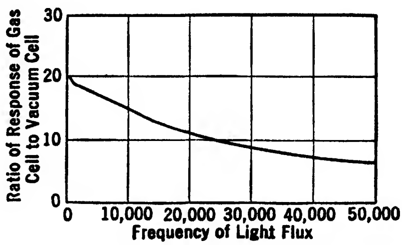


FIG. 10. Curve showing effect of dynamic response. (Courtesy Zworykin and Wilson.)

Photovoltaic Devices. A photovoltaic device is one in which the energy in incident light creates a difference in potential. In 1839 E. Becquerel discovered that, when one of two electrodes immersed in an electrolyte is illuminated, a difference in potential appears between the electrodes. This difference disappears when the electrode is in darkness. This principle has been employed in a commercial photovoltaic cell known as Rayfoto. The cell consisted of a sensitive electrode of cuprous oxide and an anode of lead immersed in an electrolyte of lead nitrate. The sensitivity of this cell was about 150 microamperes per lumen. This cell and other electrolytic types of photovoltaic cells have been displaced by the dry type of device.

In 1876, while experimenting with the conductivity of amorphous selenium rods embedded in iron, Adams and Day discovered that a difference of potential was created when light fell on their apparatus. This phenomenon of the creation of a potential difference by light falling on a junction of selenium and iron was rediscovered by Charles Fritts in 1884. Forty-six years later (1930), B. Lange observed the phenomenon for the third time.

The photovoltaic devices of the dry type consist of elements similar to those used in the blocking-layer type of rectifier—copper oxide on copper and selenium on iron. Curiously, when the device is used as a photovoltaic device the direction of the electron flow in these elements may be opposite to that observed when it is used as a rectifier. The action of a simple photovoltaic cell is illustrated in Fig. 11 (part *a*). Light falling upon a thin semiconductor layer (copper oxide or selenium) causes electrons to move to the metal (copper or iron). The elec-

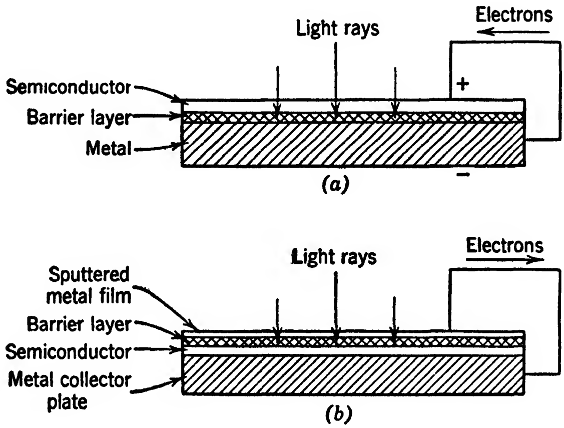


FIG. 11. Construction of photovoltaic cells: (*a*) back plate type; (*b*) front plate type.

trons that move to the metal are trapped, and they build up a negative potential in the metal. In the meantime, the semiconductor has lost electrons and assumes a positive potential, and a difference of potential is built up between the two electrodes. This difference of potential does not become very high (about 0.3 volt) because the electrons can leak back to the semiconductor through the barrier layer.

The theory of the photovoltaic cell lies in the action taking place in the barrier between the metal and the semiconductor. The semiconductor is in crystal form, and the crystals may touch the metal at points so that the area of contact is small. These crystals have a rectifying action and offer a high resistance to current flow across the barrier in one direction, but a low resistance for the opposite direction of flow. Light falling on the semiconductor penetrates it as far as the barrier crystals and there gives up the energy of the photons, which is sufficient to overcome the high resistance of the barrier and to cause electrons to move into the metal and become trapped. One may

wonder why the displaced electrons remain entrapped since the normal direction of electron movement via rectifying action is *from* metal to semiconductor. The answer to this question is to be found in Fig. 11 in the chapter on crystal and metallic rectifiers. From this figure it will be noted that, while the resistance to current in the forward direction is approximately zero for positive potentials of 1 or more volts, the resistance at low potentials of zero to 0.3 volt lies in the range of 400 to 150 ohms. Higher values may exist for cells of different dimensions and construction. Hence it is obvious that, if a low-resistance circuit or load is connected to the two electrodes of the photovoltaic

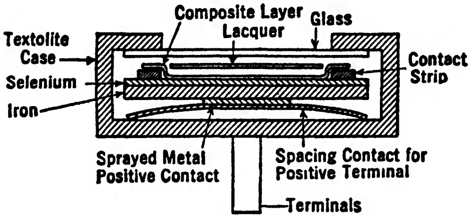


FIG. 12. Construction of a photovoltaic cell. (Courtesy General Electric Company.)

element, the leakage current can be kept small. On the contrary, a high-resistance load will result in considerable leakage.

Barrier-Layer Photovoltaic Cells. The photovoltaic devices in use today differ from the one just described in having a thin metallic layer sputtered on the surface of the semiconductor (Fig. 11*b*). Gold and platinum have been used for this surface layer. The theory of action of this cell is the same as previously explained except that the direction of electron drift is reversed. This reversal results because the active metal or metallic layer is now exposed to the light which penetrates to the semiconductor layer and gives up its energy, thus moving the electrons from the semiconductor to the surface layer. This action reverses the polarity of the cell.

It is probable that some light entering the cell may penetrate the semiconductor layer and cause a second action at the lower barrier, thus moving some electrons into the bottom collector plate. This second action will produce a counter emf effect but it will be small in magnitude.

A commercial photovoltaic cell is shown in Fig. 12. Fundamentally, the device consists of a block of iron covered by a layer of selenium. A transparent composite layer is sputtered upon the selenium surface,

after which the entire unit is covered with a film of lacquer to protect it from the atmosphere. Contacts are established with the composite layer for one electrode and with the iron back for the other electrode. When light falls on the selenium, electrons pass to the composite surface layer.

The current output of the selenium-iron device depends upon the illumination and the resistance of the external load. The variation of

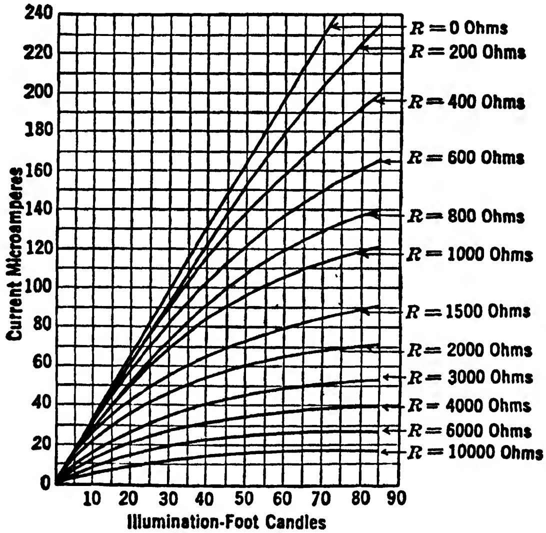


FIG. 13. Current-illumination characteristic of a photovoltaic cell for various load resistances. (Courtesy General Electric Company.)

current with these two factors is shown in Fig. 13. This variation is almost linear for zero external resistance and indicates that a low load resistance is very desirable for light-measuring devices. The change from linearity for high load resistance is due to the increase in internal leakage across the barrier. The color sensitivity of the selenium-iron cell illustrated in Fig. 12 is given in Fig. 14. The sensitivity curve compares fairly satisfactorily with that of the human eye in the range of visible light.

The characteristic of the selenium-iron cell changes with temperature, particularly where the load resistance is high. These cells should not be used at temperatures higher than 50 degrees C. The action of this device is subject to fatigue; however, the error with exposure is slight if the external load resistance is kept small. This device is also

subject to aging, but the commercial products can be aged artificially so that there is no further change in the hands of the purchaser.

The principal application of the selenium-iron photovoltaic cell has been for the measurement of light. One application known as a foot-candle meter (Fig. 15) consists of a selenium light cell and a sensitive ammeter built into a compact unit. A similar device, called an exposure meter, for photography, is illustrated in Fig. 16.

Another commercial form of selenium-barrier-type photovoltaic cell is illustrated in Fig. 17. This device is called a *photronic* cell and is supplied in three grades of sensitivity varying from 2.5 to 4.5 microamperes per foot-candle. When equipped with a viscor filter the sensitivity of these cells very closely matches that of the human eye. Photronic cells are used in circuits with sensitive d-c meters for light measurement, or they may be used to operate sensitive relays.

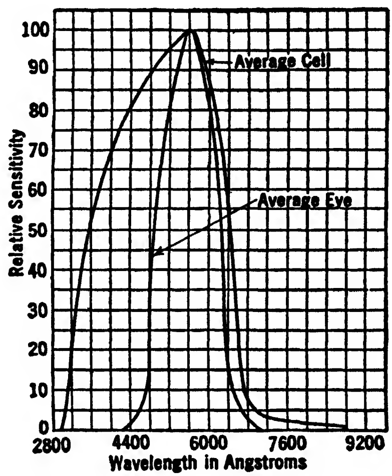


FIG. 14. Color sensitivity of a selenium photovoltaic cell.

The selenium photovoltaic cell may be connected to a meter in a simple series circuit if the resistance of the cell is suitable for damping the meter and for giving a linear deflection to the light response. If these factors are not satisfactory, shunt resistances may be employed across the meter and cell. The response of the photovoltaic cells to light change is satisfactory for light measurements, but for rapidly changing light, such as sound on film, the dynamic response is unsatisfactory.

Copper oxide photovoltaic cells have a history and a theory of operation that parallel those of the selenium-iron type. The photovoltaic action of copper oxide on copper was observed by Pfund in 1916. An English patent was obtained on the copper oxide element in 1928, and about 1935 a commercial photovoltaic cell using the copper oxide unit was placed on the American market.

A comparison of the copper oxide and the selenium-iron photovoltaic cell shows that both units are subject to effects of temperature and to aging. The important difference is in the degree of aging and the period

of aging. Here, as in the dry-type rectifier units, the copper oxide unit is subject to a greater effect of aging. Although this aging effect is not so important in a rectifier, it becomes very serious if the unit is used

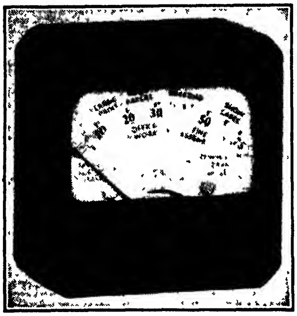


FIG. 15. Photovoltaic foot-candle meter. (Courtesy General Electric Company.)

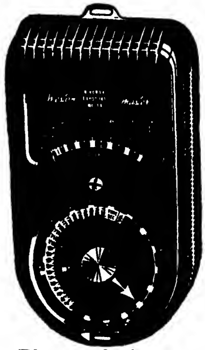
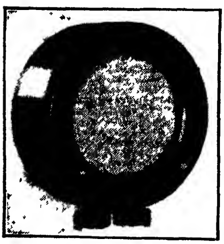


FIG. 16. Photovoltaic exposure meter. (Courtesy Weston Electrical Instrument Corporation.)

for the measurement of light. The decrease in the electron emission due to aging and to high temperatures causes the copper oxide light meter to lose its calibration, and accordingly the copper oxide photovoltaic unit has been withdrawn from the market.



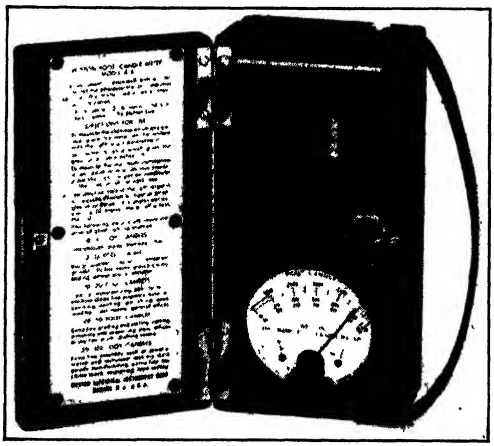


FIG. 17. Photovoltaic cell (*left*); built into a foot-candle meter (*right*). (Courtesy Weston Electrical Instrument Corporation.)

Photoconductive Devices. In 1873 Willoughby Smith observed that high-resistance elements consisting of tiny rods of selenium became better conductors when exposed to daylight or to any artificial illumi-

nation. Selenium is a chemical element which lies on the borderland between conductors and insulators. Under any condition it is a poor conductor, but in the gray crystalline form it shows a decrease in resistance under the influence of light. A large number of different forms of selenium cells have been devised to take advantage of this property of selenium. Many of these cells are old in the art so that this device antedates all other photoelectric cells in time of practical use.

In order to utilize the peculiar property of gray crystalline selenium, it must be connected as in the circuit shown in Fig. 18. When light falls on the selenium its resistance decreases and the current in the circuit rises. If R represents a marginal relay it will be operated by the increase of current due to incident light. The decrease of resistance is due to the freeing of electrons by the incident light. Apparently the energy in the photons of the incident light is imparted to some of the electrons in the outer orbits of selenium atoms so as to free them from the bonds of the nuclei. Then the potential gradient in the selenium causes these electrons to move on with greater freedom. As the electrons move on they recombine with positive ions and become neutral atoms. As long as light falls on the selenium new electrons are released, but when light ceases and recombinations are completed the electron movement falls to a low value.

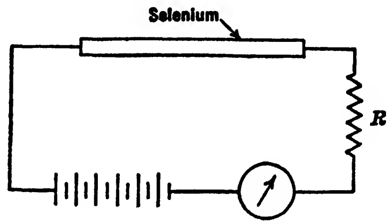


Fig. 18. Circuit for a photoconductive cell.

Since light does not penetrate the selenium very deeply and since the resistance of selenium is very high, two things are necessary in the production of a satisfactory selenium cell. First, the selenium should be used in a very thin layer to permit light penetration, and, second, the cross-sectional area for the current flow should be made as large as convenient. To accomplish these two aims many selenium cells have been made as illustrated in Fig. 19. A metal film such as gold leaf is placed on a plate of glass or other insulating material. The film is scratched by a zigzag line of grid type. Then a thin layer of selenium is placed over the whole plate and given a heat treatment to change the selenium to the gray crystalline form. After this treatment the plate may be sealed in a glass tube to protect it from the atmosphere. It will be observed that the element consists of two electrodes separated by a long, narrow bridge of selenium. When light falls on the active

side of the plate the resistance of the entire volume of selenium in the bridge is changed and a relatively large change in current may be obtained. Commercial selenium cells have shown a current ratio from light to dark as high as 4 under an illumination of 10 foot-candles.

In earlier times selenium cells were used for control operation such as turning on or off artificial light with the coming or passing of daylight. They were improved sufficiently in later years to make them satisfactory for sound reproduction from film but did not receive much recognition. If its characteristics had been improved at an earlier date, the selenium photoconductive cell might have had a wide application.

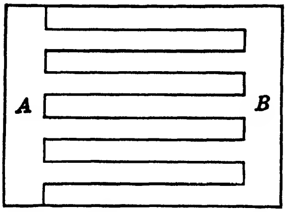


Fig. 19. Construction of a photoconductive cell.

Comparison of Photoelectric Devices. In passing, it is well to note that the photoemissive and the photoconductive devices require an outside source of d-c supply for their operation, whereas the photovoltaic device gets its energy direct from the light. The photoemissive device produces a very small current and an amplifier is needed before any useful work can be performed. The

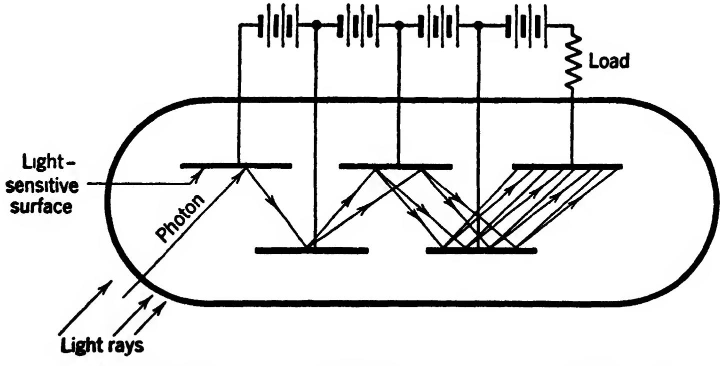


Fig. 20. Principle of a photoelectric electron-multiplier tube.

photovoltaic and photoconductive devices may operate a sensitive relay directly without any amplification. The photoemissive devices furnish a signal accurately proportional to the incident illumination and the voltage of their output is satisfactory for amplification. The photovoltaic device feeding into a low load resistance produces a current proportional to the incident illumination. When this device feeds

into a high-resistance load the voltage produced is not linear with the incident light. This fact plus the low magnitude of its voltage makes the device unsuitable for the grid supply of an amplifier.

Electron-Multiplier Tubes. Photoelectric emission is utilized for amplification of a light signal in a tube known as an electron multiplier. The theory, circuit, and one schematic type of construction of the de-

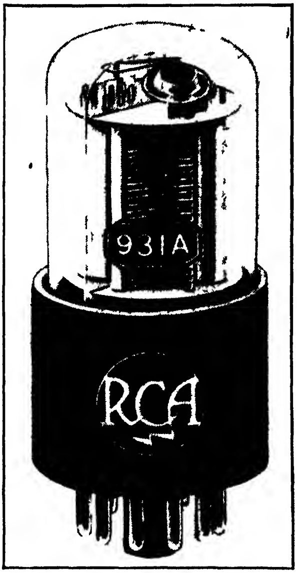


FIG. 21. Electron-multiplier tube. (Courtesy Radio Corporation of America.)

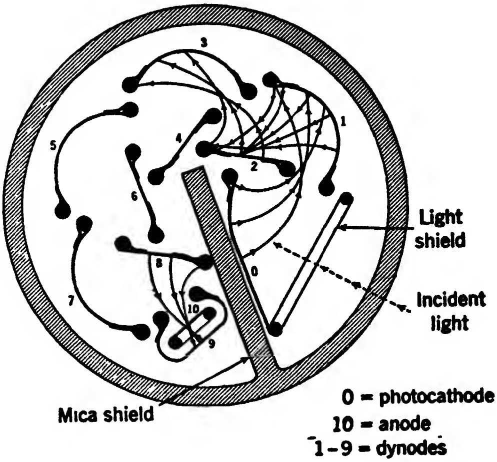


FIG. 22. Cross section of the electron-multiplier tube in Fig. 21.

vice are illustrated in Fig. 20. The first of a series of electrodes in a vacuum chamber has a light sensitive film which emits electrons in proportion to the incident light. The emitted electrons are attracted to a second electrode held at a higher positive potential sufficient to cause secondary emission. These electrons of secondary emission are drawn to a third electrode where an increased number of secondary electrons are emitted. One primary bombarding electron may emit from 1 to 10 secondary electrons, depending on the electric field and work function of the electrode surface. Thus through a series of electrodes, each having a higher potential than the preceding one, any amount of amplification may be attained theoretically. A practical limit in multiplication arises through space-charge effects and power dissipation in the final stages.

A commercial form of electron-multiplier tube is shown in Fig. 21 and an enlarged cross section of the device in Fig. 22. In this device incident light entering the tube releases electrons from the light-sensitive cathode *O* from whence they are drawn to electrode 1 where secondary emission begins and continues at increasing values throughout nine successive stages. Development engineers worked out an ingenious design for incorporating several anodes in a very compact space, the tube being only 1⅛ inches in diameter. With the use of relatively high potentials it is possible to attain an amplification of one million, though such a value is not attained in commercial usage.

The electron multiplier can be used as a combined photoelectric pick-up and amplifier for minute light variations, for sound on film, and in television.

Problems

1. The vacuum phototube of Fig. 4 is connected in the circuit of Fig. 3 with an applied d-c potential of 250 volts and a resistance of 10 megohms. Let the light falling on the cathode of the tube change from zero to 0.1 lumen and 0.5 lumen. Calculate the voltage drop across the resistor for each case. (Use curves of Fig. 7 for this tube.) Does the answer check with law No. 1 for photoelectric emission?

2. Connect the gaseous phototube of Fig. 4 into the circuit of Fig. 3 using a 90-volt battery and a 5-megohm resistor. Assume a light flux of 0, 0.1, 0.2, 0.3, 0.4, and 0.5 lumen falling on the cathode of the tube, and determine the voltage drop across the resistor for each case using the curves of Fig. 8. Plot curve of voltage drop versus lumens.

3. Repeat Problem 2 using a 1-megohm resistor. Compare the results of Problems 2 and 3. Which of the above tubes should be used for accurate light measurement?

4. Design a circuit for firing a thyratron using the increasing light falling on a phototube.

5. Refer to Figs. 1 and 5 and select a phototube surface for use with (a) ultraviolet light, (b) infrared light, and (c) incandescent lamp light.

6. Develop the following equations for the threshold frequency and threshold wavelength in terms of the work function in equivalent electron volts.

$$f = 2.415 \times 10^{14}\, \phi \text{ (work function)}$$

$$\text{angstroms} = \frac{12{,}400}{\phi \text{ (work function)}}$$

Hint: Use equation 1 of this chapter, equation 10 of Chapter I, and the value for Planck's constant

$$h = 6.624 \times 10^{-34} \text{ joule per second}$$

7. Calculate the threshold frequency for copper, nickel, and barium.
8. Calculate the threshold wavelength for sodium and cesium.

REFERENCES

ZWORYKIN, V. K., and E. D. WILSON, *Photocells and Their Application,* John Wiley & Sons, 1934.

LANGE, BRUNO, *Photo-Elements and Their Application,* Reinhold Publishing Corp., 1938.

Chapter XVI

SPECIAL PHOTO APPLICATIONS

Oscilloscope.* A cathode-ray oscilloscope is an instrument designed for the analysis of electrical circuits by a study of the wave forms of voltage and currents at various points. The instrument may be employed to study any variable, within the limits of its frequency response characteristic, that can be converted into electrical potentials. Common variables are sound, vibration, light, and all forms of variable impedances such as resistance, inductance, and capacitance.

The basic elements of an oscilloscope are a cathode-ray tube, amplifiers, and a power supply. The construction and general theory of the cathode-ray tube was covered at the end of Chapter IV and the reader may wish to review that discussion before proceeding with this chapter. The frequency limitation of an oscilloscope is determined by the electron transit time across the face of the deflection plates in the cathode-ray tube and by the functioning of the amplifier circuits. In the cathode-ray tube the transit time is of the order of 0.001 microsecond so that little error will be present in frequencies up to 100 megacycles. Amplifier circuits do have frequency limitations and they become a controlling factor in applying the oscilloscope to many problems.

The basic theory of action of the cathode-ray tube as applied to the oscilloscope may be reviewed by observing the right side of Fig. 1. The electron beam moving perpendicularly to the page passes between the parallel deflection plates HH' and VV' and forms a luminous spot on hitting the fluorescent screen. With zero potentials applied to the deflection plates and the proper adjustment of the tube circuits the luminous spot should fall at the center of the screen. If a varying potential (a-c) is placed across plates HH', the luminous spot will be caused to travel back and forth along the horizontal line hh'. If the frequency of the variation of voltage is 16 cycles or more, the hori-

* The terms oscilloscope and oscillograph are used interchangeably for this device. The contraction "scope" is frequently used.

zontal trace will appear as a solid motionless line because of the persistance of vision of the human eye. In a similar manner, if a variation of voltage is applied only across deflection plates V and V', the moving luminous spot will create a vertical line or trace as indicated by vv'. When varying potentials are placed across both pairs of deflection plates simultaneously, a picture is created on the screen which gives information regarding the wave form of the applied signals. In general, some independent or known signal is applied across the plates

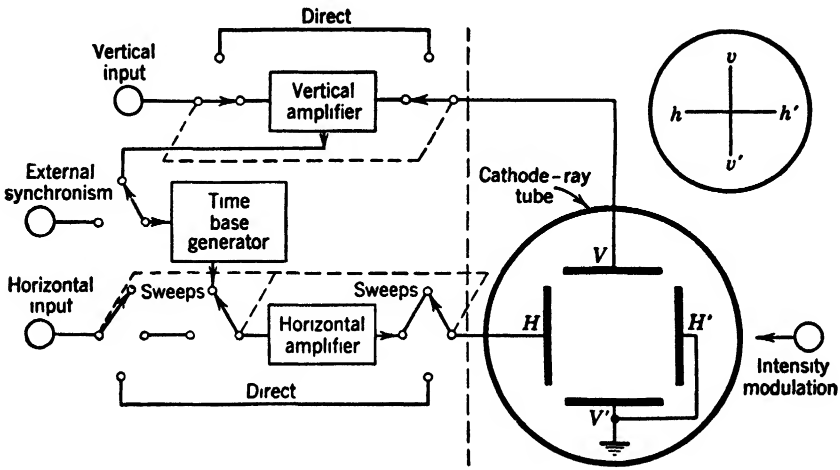


FIG. 1. Block diagram of a simple oscilloscope.

HH' while an unknown or dependent signal is applied across the plates VV'. The former is called the horizontal signal and the latter the vertical signal. For the determination of wave shapes a synchronized sawtooth signal is used for the horizontal while for the study of frequency and of phase angle a sine-wave a-c signal is used.

The block diagram of a simple oscilloscope is shown in the complete Fig. 1. The circuit is simplified and the cost of the device has been reduced by grounding H' and V'. The vertical input signal is fed between the post so marked and the ground. If this signal is of sufficient strength it is supplied directly to the deflection plates V and V', but if the magnitude is of insufficient value to produce a satisfactory deflection it may be passed through an amplifier for controlling the final signal strength applied to the plates. In a similar manner an external horizontal signal may be applied through the lower circuit group to plates direct (lower position of ganged switches) or via an amplifier

(middle switch position). For wave analysis it is generally desired to synchronize the "sweep" signal with the applied vertical input signal. This is accomplished by controlling the time base (sawtooth) generator from the vertical input with the connection as shown in Fig. 1. The latter may be considered as an internal horizontal input. For some applications it is desirable to control or modulate the intensity of the luminous spot on the screen by some external signal. This result is accomplished by the intensity modulation indicated on the right of diagram. Intensity modulation in this circuit is effected by feeding a signal into the control grid-cathode circuit of the cathode-ray tube. A better understanding of the oscilloscope may be obtained by expanding the block diagram of Fig. 1 into component schematic circuits.

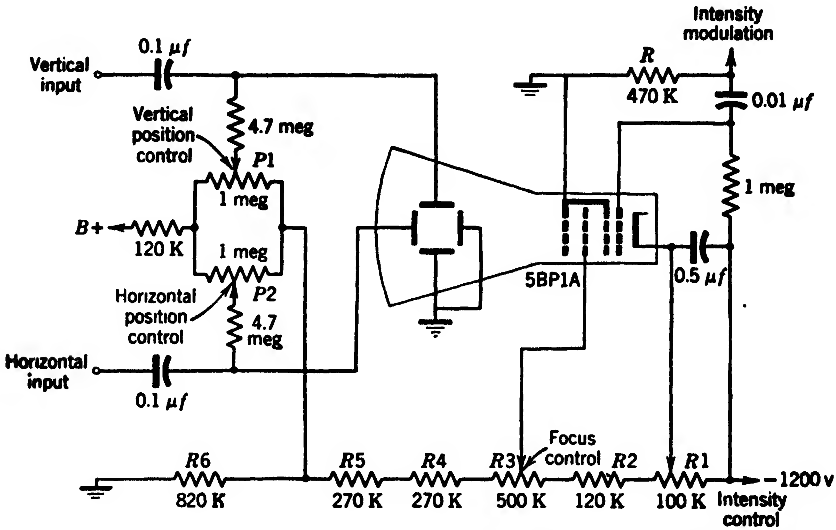


FIG. 2. Simplified schematic of cathode-ray tube circuits.

A cathode-ray tube circuit for a commercial oscilloscope, showing the magnitude of the components, is given in Fig. 2. The second anode and the deflection plates are held at or near ground potential by ground connections. The cathode and control grid are maintained at a potential of approximately −1200 volts, giving a maximum electron-accelerating potential of 1200 volts, from potential divider resistors $R1$ through $R6$. Focusing control is attained by connecting the second anode to the variable resistor $R3$. The luminous-spot intensity control is attained by varying the voltage between the cathode and the

control grid across a section of resistor $R1$. The position of the focal spot may be centered between the pairs of deflection plates by adjustment of potentiometers $P1$ and $P2$. It will be noted that the potential between the ends of the potentiometer resistances varies from negative (point between $R5$ and $R6$) to B+ (above ground).

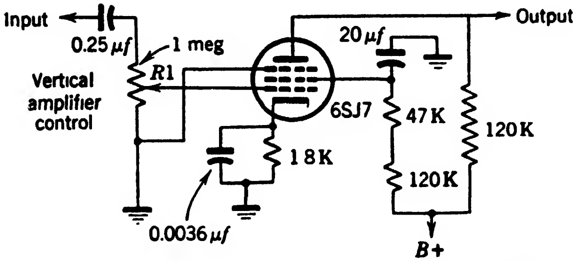


FIG. 3. Simplified circuit of an amplifier for an oscilloscope.

The amplifier circuits for the vertical and horizontal inputs to the deflection plates use pentodes as indicated in Fig. 3. The time base or sawtooth generator uses the circuit and control as indicated in Fig. 4. This is a relaxation oscillator circuit using a thyratron. The voltage supplied to cathode and anode circuit is determined by the voltage

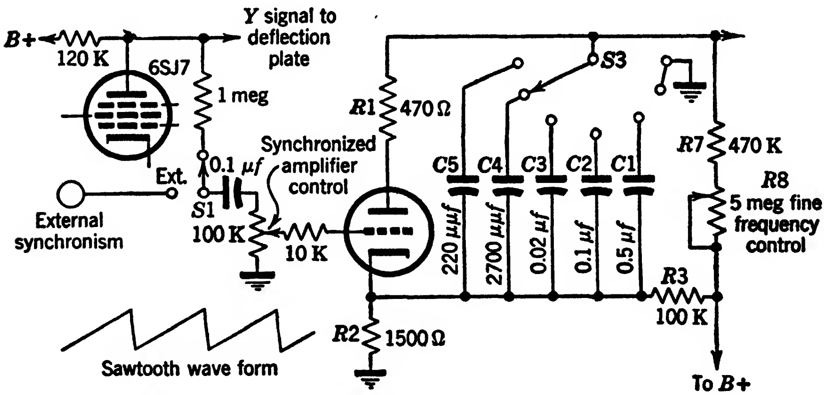


FIG. 4. Simplified schematic of a sweep circuit.

drop across the potential divider $R2$ and $R3$. For the setting shown, capacitor $C4$ charges through resistors $R8$ and $R7$ until the thyratron reaches its firing potential. At this point, $C4$ discharges rapidly, lowering the potential across the cathode-anode circuit below the conduction point. Then $C4$ recharges slowly and the process repeats, giving the

sawtooth wave form of voltage as shown in the lower left-hand corner of Fig. 4. The instant of firing (trigger action of grid) may be controlled by the connection to vertical amplifier (left) and thus give synchronous timing with the vertical input signal. Again synchronous timing from an external signal may be attained by movement of switch *S*1 to the left. An excessive synchronizing signal impairs the wave form of the sweep oscillator. Coarse frequency control can be ob-

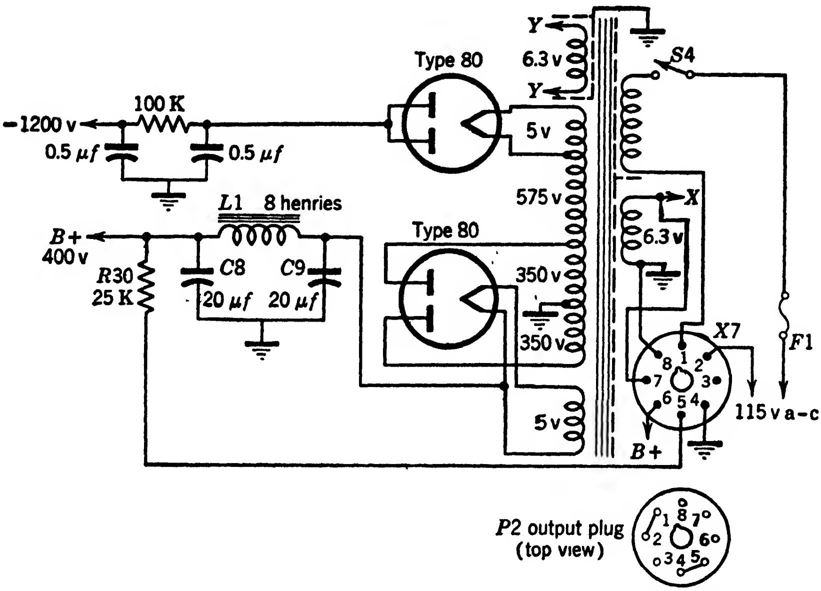


FIG. 5. Simplified circuit of the power supply for an oscilloscope.

tained by the capacitor selection of switch *S*3 and fine adjustment of frequency control can be effected by variation of resistor *R*8.

The schematic circuit for the power supply of the oscilloscope is given in Fig. 5. The top tube giving half-wave rectification furnishes the high-voltage (1200 volts) supply through a resistor and capacitor filter for the electron gun circuit. The lower tube gives full-wave rectification for a 400-volt supply for the tubes and deflection plates in the instrument plus a reserve for operating some electronic equipment accessory to the oscilloscope. The output plug in the lower right corner is for a take-off of B supply for accessory equipment and serves as a safety "disconnect" of the 115-volt, a-c supply. The safety feature results because this plug must be pulled out to remove the case from the instrument.

A commercial oscilloscope (oscillograph) which embodies the preceding circuit is illustrated in Fig. 6.

The action of the oscilloscope when studying a wave form plotted

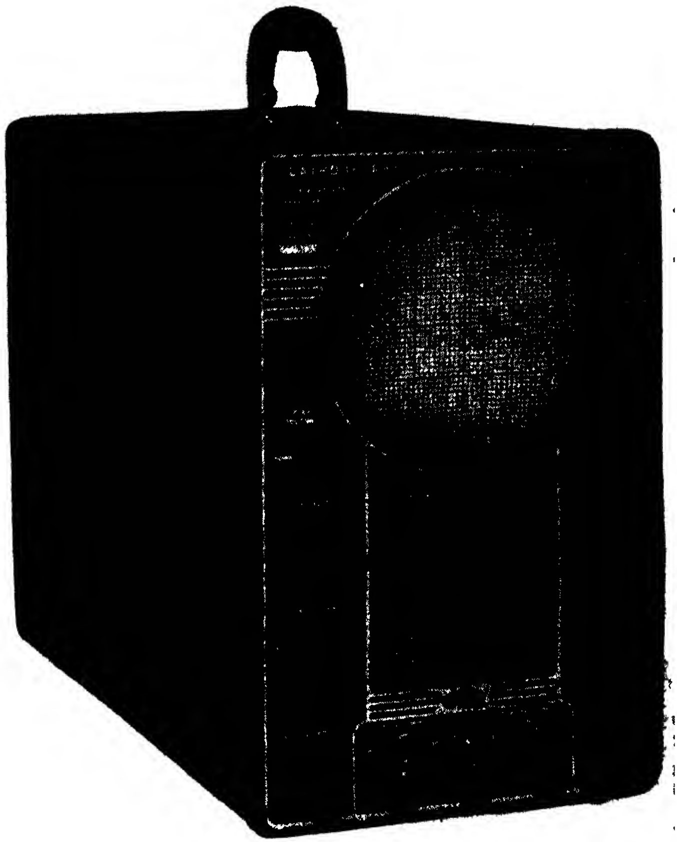


FIG. 6. A commercial oscillograph. (Courtesy Allen B. Dumont Laboratories, Inc.)

against time is illustrated in Fig. 7. A sine wave of voltage is applied to the vertical input (left) and a sawtooth wave of the same frequency is applied across the horizontal input. The magnitude of the vertical and horizontal sweep is governed by respective input signal control. For comparative and quantitative measurement of the signals a transparent screen containing cross-section lines is placed on the face of the cathode-ray tube.

When the signal inputs to both the horizontal and the vertical deflection plates are a-c voltages, the resulting action and pattern on the screen are illustrated in Fig. 8. This form of pattern is known as a Lissajous figure, named after the nineteenth-century French scientist. The Lissajous pattern may be traced by joining intersections from like numbered points on the signal. Lissajous figures are used for determining the frequency of unknown signals. Such frequency determination requires the use of one signal of known frequency which is impressed on the horizontal input. If the signals of the same frequency and magnitude and 90 degrees out of phase are impressed across the vertical and horizontal inputs, the trace will be a circle, as shown in part 1 of Fig. 9. Thus the circle becomes the pattern for a frequency ratio of 1/1. A horizontal frequency ratio of 3/1 is illustrated in Fig. 8 and several other ratios in the views of Fig. 9. The frequency relationship is determined by the ratio of the number of loops touching two mutually perpendicular sides such as AB and BC of part 5 of Fig. 9. The algebraic rule for the determination is

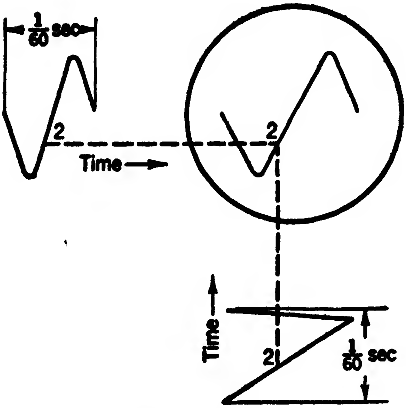


FIG. 7. Projection drawing of a sine wave applied to a vertical axis and a sawtooth wave of the same frequency applied simultaneously on a horizontal axis.

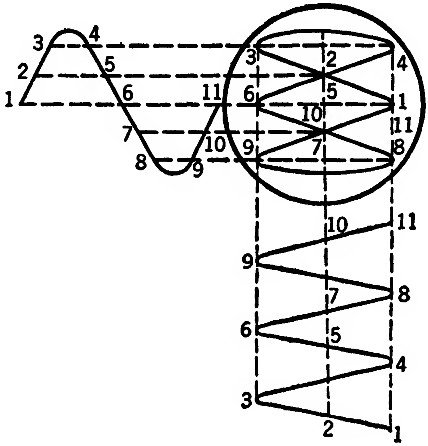


FIG. 8. Projection drawing showing the resultant Lissajous pattern when a sine wave applied to the horizontal axis is three times the frequency applied to the vertical axis.

$$\frac{\text{Frequency on horizontal axis}}{\text{Frequency on vertical axis}} = \frac{\text{Number of loops intersecting } AB}{\text{Number of loops intersecting } BC}$$

The phase-angle difference between two signals of the same frequency can be determined by the use of Lissajous figures produced by impressing these signals on the horizontal and vertical input to the oscilloscope. In making this determination it is important (1) that the luminous spot be centered on the screen of the cathode-ray tube, (2) that both horizontal and vertical amplifiers have been adjusted to give exactly the same gain, and (3) that the calibrated scale be set to coincide with the displacement of the signal along the vertical axis. The calculation for the angle of phase shift is made by use of the formula

$$\sin \theta = \frac{Y \text{ intercept}}{Y \text{ maximum}}$$

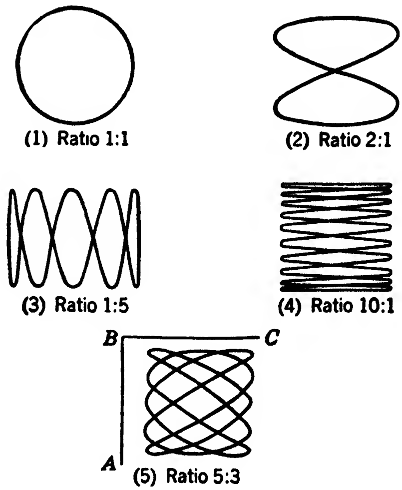


FIG. 9. Lissajous patterns.

where the Y intercept is the magnitude of the $+Y$ intercept of the trace on the Y axis and the Y maximum is the peak or maximum + value attained by the trace. Several views of patterns and the calculation of the phase-angle difference are illustrated in Fig. 10.

The frequency of a series of pulses and the duration of a single

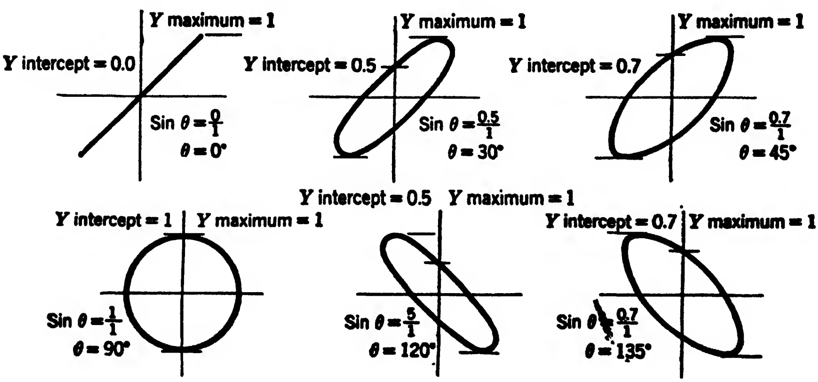


FIG. 10. Examples showing the use of the formula for determination of phase difference.

pulse may be determined by impressing the impulse signal upon a known sine wave through intensity modulation. Intensity modulation is the result of applying a signal of varying potential to the control grid of the cathode-ray tube, thereby varying the intensity of the trace at the frequency of the signal applied. The result of such intensity modulation is illustrated in the oscillogram of Fig. 11.

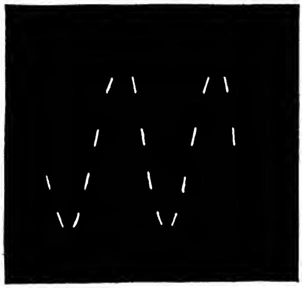


FIG. 11. Example of intensity modulation with a sine-wave signal. (Courtesy Allen B. Dumont Laboratories, Inc.)

Electron Microscope. The electron microscope is a modified and enlarged cathode-ray tube for magnifying minute objects. The optical form of microscope is limited in its effective resolving power by the wavelength of visible light to about 1000 diameters. With ultraviolet light the effective resolving power can be increased to 1500 diameters. Electrons have wave properties which permit them to be used to give resolving power fifty times as great as for light rays.

The first American commercial model of an electron microscope was designed and built by Zworykin and Morton in 1940. The principle of operation of this new device in comparison with an optical microscope is illustrated in Fig. 12. In the light microscope (*a* of the figure), light rays from a lamp are formed into a parallel beam and directed on a specimen S by the condenser lens L_1. The image of the specimen then falls on the objective lens L_2 which focuses and magnifies it, producing an enlarged image I_1. Part of this enlarged image is further magnified by the projector lens L_3. The twice-enlarged image I_2 is seen by the eye.

In the electron microscope (*b* of Fig. 12), the source of action is a hot cathode which emits electrons. Beneath the cathode there is an electron gun and anode which gives the electrons a high velocity downward. A coil or solenoid L_1 produces a magnetic field which has a focusing action and bends the paths of these electrons into a parallel beam directed on the specimen S. The electrons in the beam are affected in a varying degree, depending on the density of different parts of the specimen. Those that pass through are brought to a focus by the field of a second coil L_2 and form an enlarged image I_1. The electron rays that form a section of this image are in turn deflected and magnified by the field of a third coil L_3 and caused to form a larger image I_2. Since this image I_2 is formed by an electron beam it is not

visible. Accordingly, a fluorescent screen is placed so that the electron beam falling on it produces a visible image. If a directly viewable image is not desired, a photographic film is used in place of the fluorescent screen.

An improved commercial model of an electron microscope is shown in Fig. 13. This model is 75 inches high and 24 inches wide. It is capable of making magnifications varying from 100 up to 20,000 diameters via direct viewing. In many cases the picture (negatives) may be given a useful photographic enlargement up to 100,000 diameters. Greater enlargement does not furnish any additional or useful information. A picture taken with the electron microscope showing a total magnification of 94,000 diameters is shown in Fig. 14. The location of several of the operating units of the electron microscope is indicated on Fig. 13. The vacuum in the microscope column is created by an oil-diffusion pump. Pumping down to an operating vacuum of approximately 10^{-6} millimeter Hg requires from 1½ to 2 minutes. The specimen is placed in the vacuum chamber near the top of the column and the cover to this opening is held closed by the vacuum. Since electrons will not penetrate glass the specimen holder consists of a film of nitrocellulose 1/1,000,000 centimeter thick held on a fine wire mesh. The photographic film is inserted into the vacuum column by a gate within easy reach of the operator. The microscope column, the oil-diffusion pump, and the power supply are built into a single cabinet as shown in Fig. 13. The voltage regulator and the mechanical first-stage pump are separate and can be located in another room.

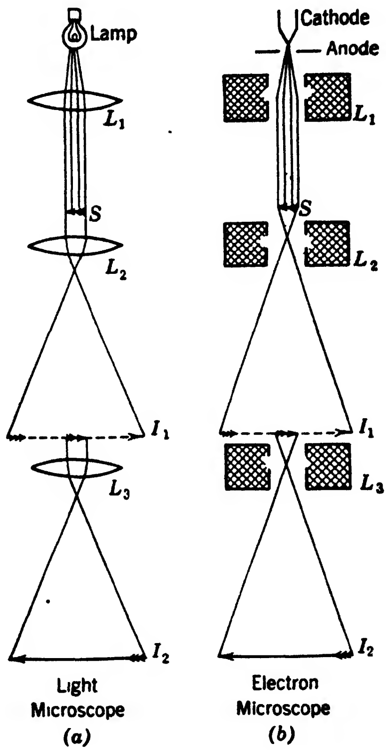


FIG. 12. Comparison of an optical and electron-microscope lens system.

The magnetic flux lenses in the electron microscope consist of iron-encased coils. The iron casing for the coils has an air gap on the inside

so shaped as to give flux paths at the center of the coil which will deflect the electron in the desired directions. A suggestion of the construction is given in Fig. 15. The magnification secured in the electron

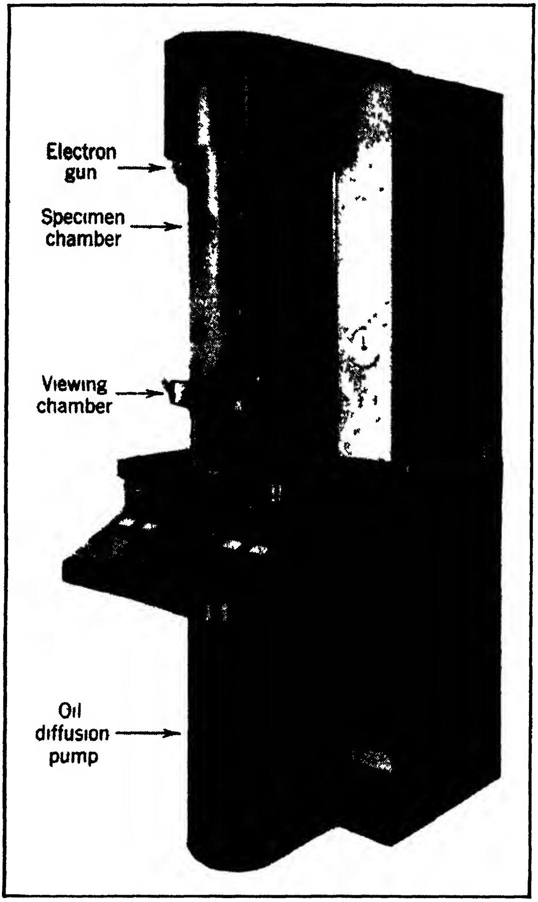


FIG. 13 Universal-type electron microscope. (Courtesy Radio Corporation of America.)

microscope can be controlled by variation of the current in the magnetic lens coils and by variation of the velocity created by the electron gun.

The electron microscope can be used as an electron diffraction camera. This application requires the removal of the specimen from the upper chamber and its insertion in the lower chamber which is in

the position of L_3 in Fig. 12. The diffraction pattern is photographed in the same manner and position as for magnified images.

The introduction of the electron microscope has opened a large field in scientific and industrial research. It has made possible the study of the structure of bacteria and virus not discernible by the optical microscope and the study of minute particles of matter in the fields of chemistry, metallurgy, botany, and food products.

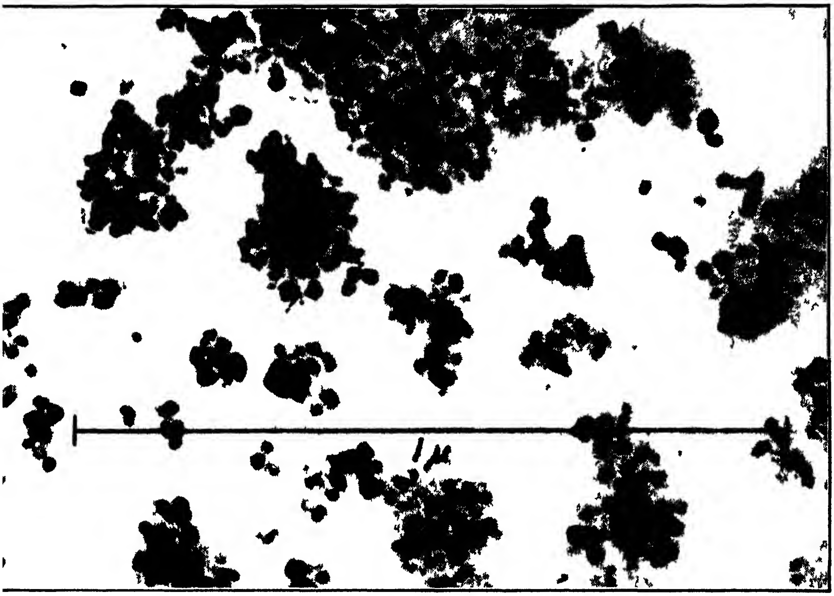


FIG. 14. Colloidal titania (titanium dioxide) photographed with the electron microscope and enlarged to 94,000 diameters. (Courtesy Radio Corporation of America.)

Stroboscope. The stroboscope is an instrument for studying or observing a periodic or varying motion by means of light periodically interrupted. Such an instrument comprises a luminous-arc discharge tube containing gas or vapor which is caused to flash by periodic transient currents. If a rotating shaft, wheel, or gear is given a flash of light once for each revolution it will appear to be at rest, or if the flashes occur at $1/n$th of its revolutions it will appear likewise to be stationary. Again, if the period of the flashes is slightly more than the period of the revolutions, the device will appear to be turning forward slowly at a speed determined by the difference in time of the flashes and period of one revolution. Similarly, a flash period less

than a revolution period will give the illusion of a slow backward motion.

This phenomenon of the stroboscope has useful applications, chief among which is the measurement of speed. When the rotating machine appears at rest the rate of flashing is the same or a submultiple of the speed, and if the flashing rate is adjustable and calibrated the stroboscope becomes a tachometer. A commercial device known as a Strobotac built for speed measurement is shown in Fig. 16. This instrument has a calibrated dial and gives an accuracy of ± 1 per cent of the dial for readings above 900 revolutions per minute when the Strobotac is standardized in terms of a frequency-controlled power line. The advantages of this type of speed measurement are that (1) no power is absorbed from the mechanism and (2) it can be used to measure the speed of machine elements inaccessible to ordinary tachometers. Thus speeds may be determined for small electric motors and delicate mechanisms which would slow down or stop if minute amounts of power were drawn from them. Likewise, the speeds of inaccessible gears or wheels in a complicated mechanism may be obtained. A second important application of the stroboscope is the slow-motion study of high-speed motion for determining vibration, tension, chattering, whip, and other irregularities that may be present in rotating and repetitive forms of motion. Thus mechanical defects and troubles may be located and remedied in many kinds of industrial machines.

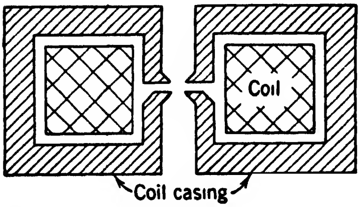


FIG. 15. Cross section of a magnetic focusing lens.


FIG. 16. A commercial form of stroboscope called a Strobotac. (Courtesy General Radio Company.)

Several different tubes and circuits may be employed in designing a stroboscope. The particular tube or circuit chosen will depend upon the amount of light desired, the period of the flash, the application, the cost, and other factors. Theoretically, any cold-cathode gaseous or vapor tube may be employed.

For low brilliance of light a small neon glow lamp (see page 258) may suffice or the firing of a thyratron gives some light. Long-column gaseous and mercury-pool and mercury-vapor tubes are often employed. The most widely used tube for stroboscopic applications is the strobotron which was shown in Fig. 37 in Chapter XII. This tube was developed at The Massachusetts Institute of Technology by Edgerton and Germeshausen. The special features of this tube are that it operates at a relatively low voltage, conducts a very high peak current of short duration, giving a very brilliant flash, and is easily controlled through two grid electrodes.

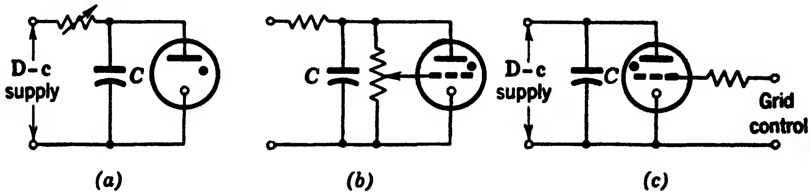


Fig. 17. Control circuits for flash bulbs.

Three simple types of circuits for the operation of a stroboscope are suggested in Fig. 17. In all these and other circuits a large capacitor C is needed to supply a heavy discharge current and to give a brilliant and quick flash. The circuit of *a* is a relaxation oscillator circuit using a two-electrode gaseous tube. The flashing period may be adjusted by the variable resistor which controls the time for the capacitor to charge up to the breakdown potential of the tube. Part *b* of Fig. 17 suggests the use of a self-exciting grid control for initiating the breakdown of the flash tube. All present-day applications use a flash tube with an internal or external grid for initiating the firing because of the relatively high breakdown potential between the cold cathode and anode. The circuit of part *c*, Fig. 17, suggests some form of external grid control. Such grid control may be produced by a mechanical contactor located on the machine which is being studied by stroboscopic light. In other cases the grid control may be effected by an auxiliary oscillator or timing circuit.

High-Speed Photography. High-speed photography is the art of taking pictures of high-speed motion by using brilliant flashes of light of a few microseconds' duration. This new art has been made possible by investigations and developments of Edgerton and Germeshausen. High-speed photography utilizes the general principles discussed under the stroboscope but requires more brilliant flashes of shorter time

duration. One commercial form of a flash lamp for high-speed photography is illustrated in Fig. 18. A long flash tube or column is made in a helical form in order to concentrate the light in a smaller area. The helical tube has a small diameter which gives a quick deionization time and short flash. A typical circuit for use in connection with this tube is given in Fig. 19. This circuit embodies a principle frequently employed in stroboscopes for initiating the discharge of the flash tube. This principle utilizes an ignition type of transformer for impressing a high transient voltage across the grid. In the lower right-hand corner of Fig. 19 a capacitor $C3$ is charged to the potential across resistor $R2$. The closing of switch $S1$ permits $C3$ to discharge through the primary of transformer $T1$ which, in turn, induces a high voltage (approximately 10,000 volts) between the grid and the cold cathode of the flash tube. This transient initiates cold-cathode emission and capacitor $C2$ discharges through the flash tube. Flash tubes are now employed in conjunction with portable auxiliary equipment by the press and commercial photographers for on-the-spot photography in the place of photoflash bulbs. A commercial device of this type is illustrated in Fig. 20.

FIG. 18. Flash tube for high-speed photography. Typical operating voltage, 2250 volts; flashing rate, 6 per minute; light output peak, 12,000,000 lumens. (Courtesy Sylvania Electric Products, Inc.)

High-speed photography is employed in research to study the conditions taking place in explosions, destruction of equipment by projectiles, projectiles in flight, and other transient phenomena. In some cases a rapid series of flashes is used to show changes and to determine the velocity of projectiles and other moving objects. An excellent photograph of this type is shown in Fig. 21.

Iconoscope. The iconoscope is a special form of cathode-ray tube used in television transmission for "picking up" a scene and converting it to an electrical signal. The principal parts of an iconoscope are the mosaic, signal plate, collector, and electron gun. The position of these parts in the tube is illustrated in Fig. 22. The mosaic consists of a large number of small photosensitive globules deposited on one face of a thin sheet of insulation. The globules are spaced a very small dis-

tance apart on the sheet so as to be insulated from each other. On the opposite face of the insulating sheet is a conductive film, the signal plate. Because the insulating sheet is thin, there is considerable

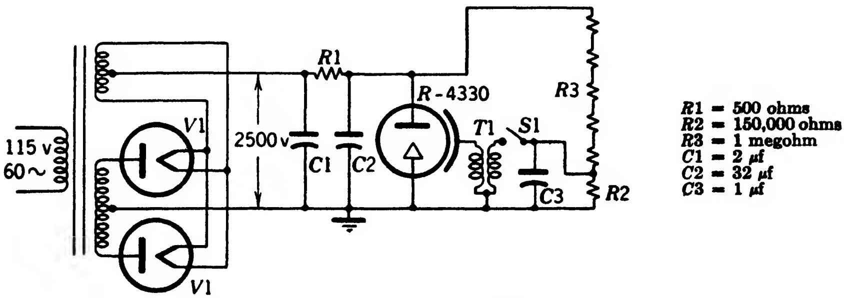


FIG. 19. Circuit for a photoflash operation.

capacitance between the globules and the signal plate. The collector is a conductive coating on the inner surface of the tube walls which collects electrons emitted by the mosaic.

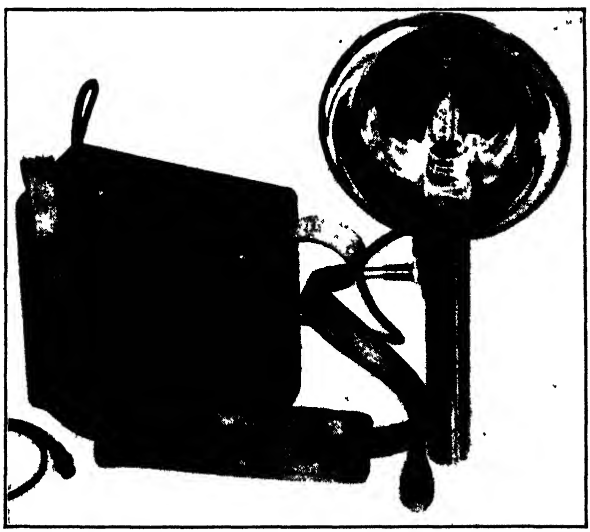


FIG. 20. Portable photoflash equipment for photography. (Courtesy Sylvania Electric Products, Inc.)

In operation of the iconoscope, an image of a scene is focused by a lens system on the mosaic, and the beam of electrons provided by the electron gun is made to scan the image. As the beam moves over the

image, there is generated at the signal plate a voltage whose magnitude at any instant depends on the image brightness at the point where the beam is at that instant. This voltage can be used as the video signal for television transmission of the scene viewed by the iconoscope. The process by which the iconoscope generates this voltage can be described briefly as follows.

FIG. 21. Multiflash photograph showing the pattern made by a golf club as it moves through the stroke. (Courtesy H. E. Edgerton.)

Consider first the action of the tube when the mosaic is scanned by the beam with no illumination on the mosaic. When the electron beam strikes a globule, the globule emits secondary electrons, the number of secondaries being several times larger than the number of beam electrons striking the globule. Some of these secondaries return almost immediately to the globule, the rest escape and go either to the collector or to other parts of the mosaic. Because the globule is insulated, its potential will change in the positive direction if the number of electrons escaping from it is greater than the number of electrons

flowing to it. The number of electrons that escape depends on the potential of the globule, the number being less, of course, the more positive the globule is. Hence, if the beam is on the globule a sufficiently long time, the globule will be driven to a positive potential at which the number of electrons escaping is equal to the number of electrons arriving. In the usual operation of the iconoscope, the time required for the beam to pass over a globule is long enough for the globule to attain this potential. The value of this potential for typical operating conditions is about 3 volts positive with respect to the collector.

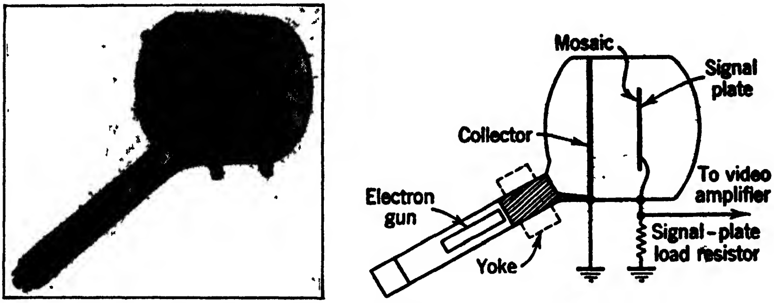


FIG. 22. Iconoscope. (Courtesy Radio Corporation of America.)

After the beam passes the globule, some of the secondary electrons emitted from the rest of the mosaic fall on the globule. The arrival of these electrons changes the globule potential in the negative direction to a new value. In a typical operating condition, this value is about 1½ volts negative with respect to the collector. With no light on the globule, the globule stays at this negative potential until the next time the beam strikes, when the globule again releases electrons and rises to its maximum positive potential of approximately +3 volts.

Consider now the action of the tube when the mosaic is scanned with part of it illuminated. During the time between contacts with the beam, both an illuminated globule and an unilluminated one receive electrons from the rest of the mosaic. Both globules, therefore, charge in the negative direction during this time. The illuminated globule, however, at the same time emits electrons, the emission being caused by the light on this globule. The illuminated globule, therefore, does not fall to as negative a potential as the unilluminated one does. Hence the next time the beam strikes, the illuminated globule does not have so far to rise to reach +3 volts. As a result, less charge is released to the collector when the beam strikes the illuminated globule than

when the beam strikes the unilluminated one. The difference in charge is approximately proportional to the difference in illumination.

Now consider the action of the tube when an image is projected on the mosaic and scanned by the beam. As the beam moves over the mosaic, varying amounts of charge flow from the mosaic to the collector, the amount of charge flowing at any instant being a measure of the light on the globules where the beam is at that instant. In other words, a video signal current flows between the mosaic and the collector. It can be seen that, since the beam current to the mosaic is constant, the video signal current must complete its circuit path through the signal-plate load resistor and the capacitance between signal plate and mosaic. The voltage developed across the load resistor by this signal current is the video signal output of the iconoscope.

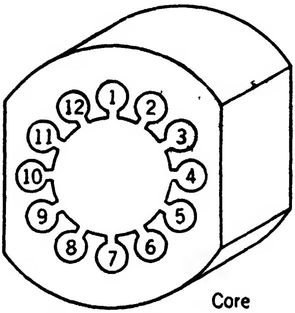


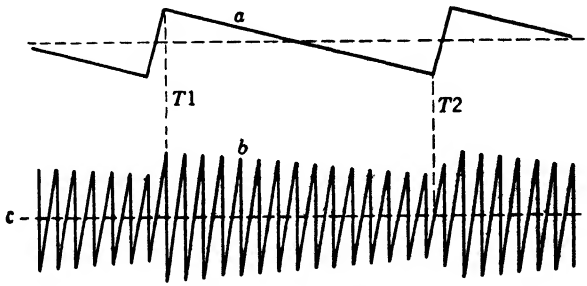
FIG. 23. Magnetic core for a yoke.

It also can be seen that, when the beam moves from a dark portion of the image to a brighter portion, the electron current from the mosaic to the collector decreases. The output voltage, therefore, changes in the negative direction. Hence the signal output of the iconoscope is of negative polarity: a highlight in the image is represented by a negative value of signal voltage; a shadow in the image is represented by a positive value.

The focusing and deflection of the electron beam is produced by magnetic fields. The focusing is secured by a solenoid coil which surrounds the neck of the iconoscope. Two pairs of coils for producing the magnetic deflection are wound in the slots of an iron core (Fig. 23). The focusing and deflection coils and their core are assembled into a single unit called the yoke which slips over the neck of the iconoscope as shown in Fig. 22.

The horizontal and vertical deflection of the electron beam is produced by sawtooth signals having wave forms as indicated in Fig. 24. The horizontal sawtooth signal causes the beam to move rapidly from left to right across the image on the mosaic and then return in near zero time to the left. While the horizontal sweep signal is causing the beam to scan the image in the horizontal direction, the vertical sawtooth signal of a lower frequency causes the beam to move slowly from the top to the bottom of the mosaic and then return with a nearly instantaneous sweep to the top. The result of these combined sweep

signals is to cause the beam of electrons to scan a rectangular area of the image focused on the mosaic as indicated in Fig. 25. This entire scanning process takes place in a period of approximately $\frac{1}{30}$ second. The ratio of the frequency of the horizontal to the vertical signal is about 500/1, which means that 500 horizontal lines or sweeps take place in $\frac{1}{30}$ of a second. During any one of these horizontal sweeps



a = vertical frequency modulating voltage
b = voltage across condenser $C3$
c = average value of voltage across $C3$

FIG. 24. Sweep signals for television.

the light intensity viewed on the mosaic may change from light to dark in $\frac{1}{500}$ of the distance along the sweep. This means that the period of a maximum light change may be

$$\frac{1}{500} \times \frac{1}{500} \times \frac{1}{30} = 1.33 \times 10^{-7} \text{ second}$$

which corresponds to a frequency of 3,750,000 cycles. Thus the *video* or image signal picked up in the iconoscope is of the order of several megacycles.

Television. Television is the art of transmitting electrically the image of an object in a fraction of a second. Such transmission may be via wire (coaxial cable) or radio. The essential elements of the television process are some form of electronic camera such as an iconoscope and a cathode receiving tube. Successful transmission of pictures via television requires (1) perfect synchronizing of the sawtooth sweep signals in the iconoscope and in the cathode-ray tube receiver and (2) the distortionless transmission of the video signal picked in the camera to the intensity control grid of the cathode-ray tube. If at every instant of time the electron beam in the receiving tube falls upon a spot on its fluorescent screen corresponding to the spot being hit by the electron beam in the transmitting camera and if the intensity of the

receiver beam corresponds to the intensity of the signal being picked up at that instant, the received image should correspond to the image in the transmitting camera.

Television has many commercial, industrial, and military applications in addition to its natural field in entertainment. Television receivers on several floors of a department store may show a style show taking place in another part of that store. The factory manager may

FIG. 25. Picture scanning in television.

view processes taking place in various places in that factory. Dangerous processes in the manufacture of explosives may be viewed through television by operators located in safe surroundings at a distant point. Guided missiles and pilotless planes may be followed and controlled through the use of television.

REFERENCES

ZWORYKIN, MORTON, RAMBERG, HILLIER, and VANCE, *Electron Optics and the Electron Microscope*, John Wiley & Sons, 1945.

GERMESHAUSEN, K. J., and H. E. EDGERTON, *Elec. Eng.*, **55**, 709 (1936).

ZWORYKIN, V. K., "Iconoscopes and Kinescopes in Television," *RCA Review*, July 1936.

Chapter XVII

SPECIAL TUBES AND CIRCUITS

X-Ray Tube. The X-ray tube is a two-electrode tube invented by Roentgen in 1895. This tube had the ability of producing electromagnetic radiations known as roentgen rays. In 1913 Doctor Coolidge made improvements which have resulted in wide and valuable applications of the tube. All modern X-ray tubes use a heated tungsten cathode partially surrounded by a cylinder or focusing cap for directing the electron emission toward the anode. The anode, generally called the target, is made of tungsten or of a tungsten insert in a copper base. The two electrodes are encased in a highly evacuated Pyrex glass tube as illustrated schematically in Fig. 1. The face of the target is at an angle to the axis of the tube so as to direct the emanating rays in a direction where they may be utilized.

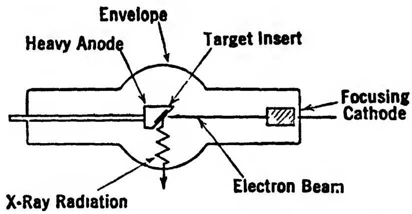


FIG. 1. Construction of an X-ray tube.

The electrons emitted by the cathode are attracted to the anode by a high potential of the order of 10,000 to 1,000,000 volts and sometimes higher. The tremendous impact of the electrons in hitting the target excites the electrons in the atoms of the target so that they send out electromagnetic radiations known as X rays. These emanations are not electrons or particles but wave energy similar to light waves. It will be recalled that the impact of electrons on atoms of gas produces an excitation of the electrons which results in the production of light. In the X-ray tube the impact of the electron is much greater because of the high potential gradients and it acts upon electrons in a solid so that the resulting waves are much shorter (higher frequency) and contain a greater energy. It is probable that the impinging electrons in the X-ray tube penetrate into the inner electron rings and even the

nuclei of the atoms constituting the target. The position of the X rays in the electromagnetic spectrum was shown on page 365. The X rays emitted by the tungsten target are a continuous spectrum of electromagnetic radiation. The minimum wavelength is determined by the peak voltage across the X-ray tube according to the equation $Vl = 12.354$, where V is the voltage across the tube measured in kilovolts and l is the minimum wavelength in angstrom units (10^{-8} centimeter). X rays are propagated with the same speed as light, follow the inverse square law, are unaffected by electric and magnetic fields, and can be reflected, diffracted, and polarized.

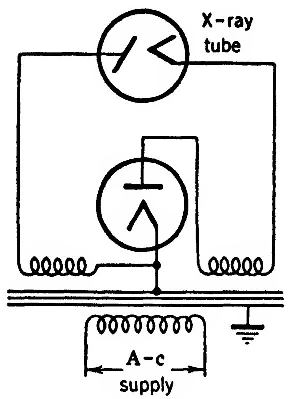


FIG. 2. Half-wave rectifier for an X-ray tube.

X rays have the power of penetration through layers of solids that are opaque to light. These X rays are invisible to the eye but can be detected after passing through solids by their action on a fluorescent screen or on a photographic film. Their ability to blacken sensitized film in a manner proportional to their intensity has been the basis of radiography. Furthermore, their ability to penetrate opaque objects is dependent upon the atomic weight (density) of the object. The penetrating power of X rays is referred to as a degree of hardness; soft rays have small penetrating power, hard rays deep penetrating power. Hardness is proportional to the frequency of radiation and depends on the voltage applied across the cathode-anode circuit. The penetrating power varies approximately as the square of the voltage. Thus in the operation of X-ray devices the hardness is controlled by varying the cathode-anode voltage, the intensity or quantity of X rays is controlled by the voltage across or the current passed through the cathode for heating, and the quantitative effect of the X rays is controlled by the time of exposure.

X-Ray Equipment. An X-ray machine consists of a high-voltage power supply, an X-ray tube, and control equipment. Since X-ray tubes are rectifiers, they may be energized by alternating current, by half-wave or full-wave a-c rectifiers, or by direct current. Since a heavy load may raise the temperature of the tungsten target to a point where appreciable electron emission may occur, there are definite limits on the use of half-wave and a-c excitation if inverse current

flow and tube destruction are to be avoided. A conventional circuit for half-wave rectification is shown in Fig. 2. It should be noted that the kenotron rectifying tube and the mid-point of the high-voltage secondary of the transformer are grounded. A second circuit for half-wave rectification is given in Fig. 3. The auto transformer on the left provides for a control of the magnitude of the peak applied voltage.

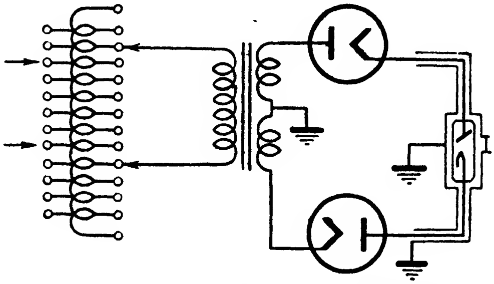


FIG. 3. Two-tube half-wave rectifier with a shockproof X-ray tube.

The grounded mid-point of the high-voltage transformer winding limits the voltage to ground at any point in the circuit to one-half of the a-c peak. The use of two kenotrons reduces the inverse peak per tube and prevents any inverse electron flow through the X-ray tube. Full-wave rectification for the X-ray tube may be secured by the use of a transformer with mid-tap and two kenotron rectifier tubes or by the use of the bridge circuit with four kenotrons as covered in Chapter XIV on power rectification and inversion. The maximum cathode-anode voltage across the X-ray tube for the preceding circuits for power supply is the a-c secondary high-voltage peak. A special power-supply circuit which applies two times the peak voltage across the X-ray tube is shown in Fig. 4. This is known as the Villard circuit and the explanation of its operation is left to the reader.

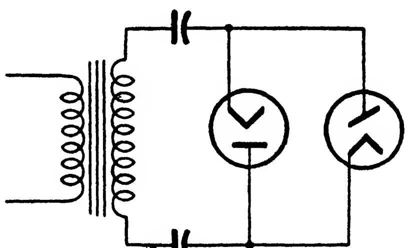


FIG. 4. Villard circuit for X-ray equipment.

Since X-ray tubes are operated at the saturation current, the electron current through the tube is determined by the size and temperature of the tungsten filament. The tungsten anode or target may be mounted on the end of a molybdenum rod or set in a copper anode.

Since only a fraction of one per cent of the electrical power put into the X-ray tube is converted into X rays, nearly all the remainder appears as heat in the target. Therefore, effective means of cooling the target must be provided. Those anodes that consist only of tungsten mounted on molybdenum rods are cooled by radiation. Those targets that are cast in copper rods are cooled by heat conduction along the copper rod to a radiator which permits the heat to be carried away by air or by oil. The anodes of some tubes which are operated continuously at high loads are hollowed out behind the target and cooled by circulating water or oil. The minimum length of an X-ray tube is determined by the spark-over distance for the maximum voltage applied to the tube. Many modern X-ray tubes are made shockproof by being operated inside grounded metal tanks filled with oil. These tubes are shorter than those operated in the open air. A shockproof tube is indicated in Fig. 3.

Applications of X Rays in Medicine. X rays are used in the field of medicine for radiography and therapy. From a knowledge of anatomy and the relative opacities of the various organs of the human

FIG. 5. X-ray photographs of human teeth: *left,* impacted tooth and fillings; *center,* abscessed tooth; *right,* permanent teeth pushing up baby teeth. (Courtesy Dr. C. J. Buster.)

body to X rays, it is possible to recognize injuries or diseases by examining the radiograph (shadows upon a photographic film). Broken or dislocated bones, defective teeth, and the presence of bullets or other foreign material are shown by the radiograph. Examples of radiographs of teeth are given in Fig. 5. Lung tissue destroyed by tuberculosis, gallstones, and ulcers of the stomach may be detected on a radiograph by the radiographic specialist. There is an optimum voltage for radiographing each part of the human body, depending on the thickness and density of the part viewed. The maximum voltage required for radiographing the human body is 100,000 volts. X rays are used to treat or destroy malignant tumors such as cancer. For

treating cancers near the surface, tube voltages varying from 10,000 to 140,000 volts are satisfactory, but for deep-seated tumors very hard X rays produced by tubes operating at 200,000 to 1,000,000 volts are required. Since these high voltages also produce soft rays which would burn the surface tissue, these softer rays are filtered out by sheets of copper and aluminum. Radiographic tubes are operated intermittently for comparatively short periods of time, whereas tubes for therapy must be capable of carrying currents of 10 to 25 milliamperes continuously. A deep-therapy X-ray tube is illustrated in Fig. 6.

Fig. 6. X-ray tube for deep therapy. (Courtesy Westinghouse Electric Corporation.)

Industrial Applications of X Rays. In industry X rays are used for fluoroscopy, X-ray diffraction, and radiography of machine parts. *Fluoroscopy* is the visual representation of the construction of opaque objects on a fluorescent screen by the transmission of X rays. This process has wide application in the inspection of foods and packaged devices before distribution. Under small-scale methods the parts to be inspected may be placed under the fluorescent screen manually whereas under large-scale processes the parts to be inspected move on a conveyer belt so that an operator makes a visual inspection as the shadowed view passes on the fluorescent screen. Tube voltages somewhat under 100,000 volts are used in fluoroscopy.

The nature and behavior of most substances depends upon the arrangement of atoms and molecules in the crystalline structure. This is true whether the material is a deep drawing steel, a bearing metal, or a tempered steel spring. Every substance such as rolled, forged, or heat-treated metal has a distinct atomic arrangement which determines its properties. Each atomic arrangement controls the diffractive effect produced by X rays. Thus, if a beam of X rays is passed through a crystal, the beam will be bent or redirected in a series of emergent rays whose separation and intensities are characteristic of the material. A radiograph of the diffracted rays constitutes a "fingerprint" of the substance because no two substances have been found to produce identical diffraction patterns.

An important application of X-ray diffraction methods is the determination of the optimum angle for cutting quartz to produce wafers (crystals) for controlling the frequency of crystal oscillators. The orientation or angle requirement for crystals manufactured before the war was met by a trial-and-error performance-selection method. During the war more crystals were required per day of manufacture than in an entire year before. Also it was necessary to produce crystals that would give the same frequency regardless of temperature.

These requirements were met by the development of an X-ray goniometer which indicates the intensity of X-ray diffraction in the direction characteristic of a certain set of planes. A sample crystal wafer is turned in the X-ray beam until this set of planes diffracts with maximum intensity. At this position the orientation of the diffracting planes is determined and the proper angle for cutting the mother quartz is determined.

Closely allied to X-ray diffraction methods is the process of microradiography. Here radiographs are made of exceedingly thin sections cf materials. These radiographs can be greatly enlarged to provide a means of adding depth to the studies of surface conditions with which microphotography is concerned. Radiographs can be applied to solutions so that the process has applications in chemistry as well as in metallurgy. These applications require tube voltages of the order of 6000 to 9000 volts.

X-ray inspection with radiography of castings, machine parts, welding structure, and intricate machinery assemblies provides the design engineer and the production department with a tool of incalculable value. Through the use of X-ray inspection, casting and electric welding techniques may be determined for producing stronger and flawless castings and electric welds. Inspection of castings before machining permits the elimination of parts that would result in inferior or rejected products. X-ray inspection permits a complete study of mechanical parts that are ordinarily obscured. It is important to observe that this method is a nondestructive inspection.

During World War II two new developments were made that have greatly extended the horizon on the application of X radiation. These developments are the 1- and 2-million-volt X-ray machines and the betatron. These devices have increased the hardness and power of penetration of X radiation and have reduced the required time of exposure for radiography.

The Million-Volt X-Ray Tubes. The new million-volt X-ray industrial units involve a new principle of action, a new type of tube, a new

type of transformer, and a new form of assembly. The relative size of the new X-ray tubes and their construction are illustrated in Fig. 7. These tubes have a filamentary tungsten cathode, a copper-backed tungsten target at the lower end of the extension chamber, and cylin-

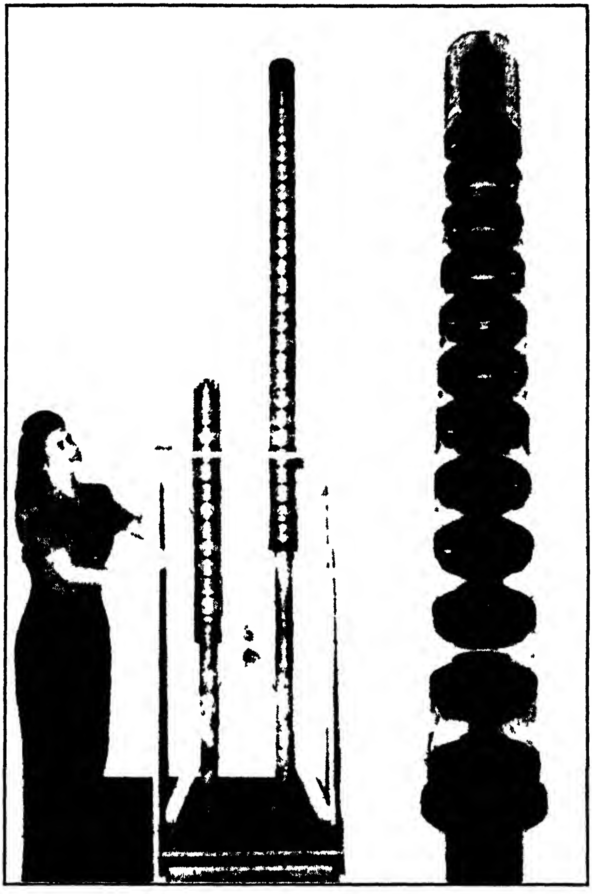


FIG. 7. Construction and comparison of the size of 1- and 2-million-volt X-ray tubes (Courtesy General Electric X-Ray Corporation.)

drical accelerating electrodes in each of the 12 or 24 intermediate sections. The accelerating electrodes serve to distribute the high potential along the length of the tube. The inside walls of the tube are sand blasted to eliminate dangerous field current which might result from the application of increased voltage to each tube section.

Adequate protection must be given to all human beings who work near X-ray equipment. This protection must cover the hazards due

to (1) high voltage and (2) the harmful effects of X radiation. The most effective bar to X radiation is sheet lead and it is often used to surround all parts of the X-ray tube except a window for the radiation. It is generally used for doors into X-ray exposure chambers and sometimes it is used to surround the entire room where X rays are used. A special form of plaster for the walls of the X-ray chamber is often used because of its lower cost. Also thick concrete walls are fairly effective for screening X rays. The advice of experts in the X-radiation field should always be secured before installation of X-ray equipment.

The Betatron. The betatron is an induction electron accelerator wherein electrons are accelerated in a magnetic field instead of by direct application of a high potential. The basic idea of accelerating electrons by magnetic induction was patented by J. Slepian in 1927. In the years following this announcement Wideroe, Walton, and Steenbeck developed equations giving the necessary conditions for an electron accelerator but were not able to produce a machine that would work. It remained for Dr. D. W. Kerst of the Physics Department of the University of Illinois to invent a satisfactory working model of the betatron (announced in 1940). Kerst assisted in the design and development of a 20-million-electron-volt (mev) commercial betatron in the early forties. A larger 100-mev commercial betatron was completed in 1945.

The basic principle of the betatron is relatively simple. In part *a* of Fig. 8 a circular iron core is surrounded by a single turn of wire *t*. If a changing flux is made to pass through the iron an emf will be induced in the turn of wire and, if the wire circuit is closed, a circulating electron current will result. Since a current is a movement of electrons, it is obvious that, if electrons are released in the path shown by *t* while the flux changes in the core, these electrons will be induced to travel around the core in a curved path and will be accelerated as long as the flux is changing in the same direction. One problem in applying this simple principle is to make the electrons follow a constant circular path. The changing flux for the betatron is produced in a three-legged transformer as shown in *b* of Fig. 8. A "donut"-shaped accelerating tube is placed around the central leg of the transformer in the plane of *de*. Exciting coils are placed on the top and bottom sections of the central leg. The transformer is excited by a resonance type of circuit as indicated in part *c* of Fig. 8. The resonant circuit on the transformer may be inductively coupled as shown, or a capacitive type of coupler may be employed.

Some of the details of the operation of the betatron are indicated in Fig. 9. The heart of the device is a torus or "donut" of glass which surrounds the changing stream of magnetic flux. The cathode and target are combined into a single unit. A heated tungsten cathode supplies electrons for an electrostatic gun which fires a stream of electrons tangentially into the central circular path. After arriving in

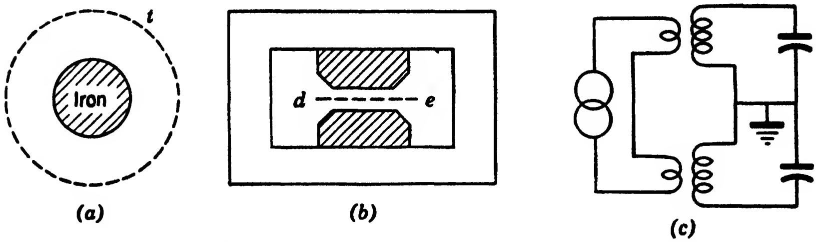


FIG. 8. Schematic construction and circuit of a betatron.

this path, the changing magnetic induction accelerates the flying electron in a circular path. When the desired velocity of the electron has been acquired, a transient current passed through a one-turn coil near each pole face having a diameter slightly less than the optimum orbit serves to force the electron into a circular path of larger diameter

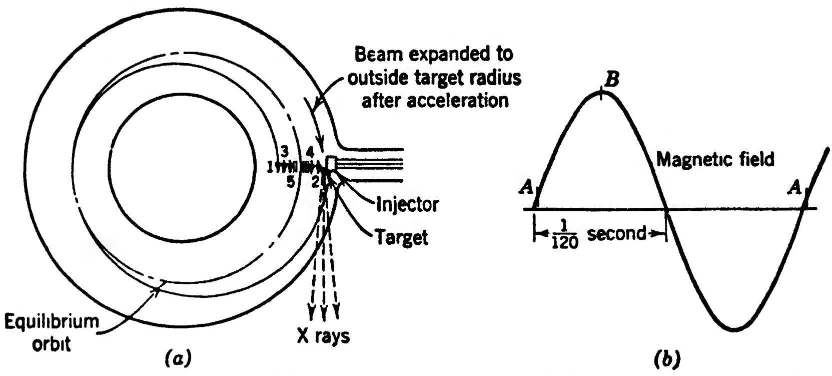


FIG. 9. (a) Schematic construction of the X-ray torus in a betatron. (b) Sine-wave magnetic field.

(beam expanded) so that it hits the target and produces X rays. The flux change in the transformer follows the sine wave shown at the right of Fig. 9. In operation the electron beam is injected at point A (near zero) on the sine curve and the beam is deflected or expanded at any point later (up to B), depending on the desired energy in million-electron-volts.

The problem of maintaining the accelerating electrons in an optimum circular path until they hit the target has been suggested. It is apparent that as the electron stream whirls and accelerates in a curved path the momentum of the electrons increases and they tend to move in a spiral path which would soon carry them outside the "donut." An electron moving at right angles to a magnetic field is acted upon by a constant force which causes it to move in a curved path (circle for constant field H). In the present problem the force of the stray field in which the electron stream moves tends to hold the path circular but, since the electron velocity and momentum is constantly increasing, the circular path will be a spiral (increasing diameter) if the field H is constant. Therefore, in order to hold the diameter of the electron path constant, it is necessary to have H increase at the same rate as the momentum of the electrons. This necessary condition has been secured by the proper shape of the iron pole faces and magnetic disks in the air gap giving a stray field flux of the necessary value along the optimum circular path of the electrons.

A 20-mev commercial betatron is illustrated in Fig. 10. The design of the betatron makes possible acceleration of electrons to very high energies with low-voltage equipment. In the 20-mev unit electrons are given the same energy they would receive if they were accelerated by a potential difference of 20 million volts, while the voltage required to create the necessary field for their acceleration by magnetic induction is about 100 volts per turn on the magnet coil. The maximum intensity of the 20-mev X rays is in the direction of the electron stream as it strikes the target. This intensity falls off 50 per cent at 4½ degrees from the center and somewhat more gradually at larger angles. The focal spot of the betatron is very small, about 0.05 inch high by less than 0.005 inch wide. This small focal spot accounts for a sharpness in detail of radiographic work when the betatron is used.

A 100-mev betatron was completed by the General Electric Company in 1945. This unit weighs 130 tons and is about 9 feet high, 6 feet wide, and 15 feet long, and it uses a 24,000-kilovolt-ampere capacitor for power factor correction. The X-ray tube is made of 16 molded glass sections connected end to end. The transformer magnet is operated on 60-cycle current with a flux density at the orbit of 4000 gauss. The electrons are injected with a voltage ranging from 30 to 70 kilovolts and, if allowed to remain in the 66-inch-diameter circular orbit for the entire quarter-cycle, they circle the magnetic flux 250,000 times, acquiring on each revolution an average additional energy of about 400 electron-volts. Thus in a period of $\frac{1}{240}$ of a

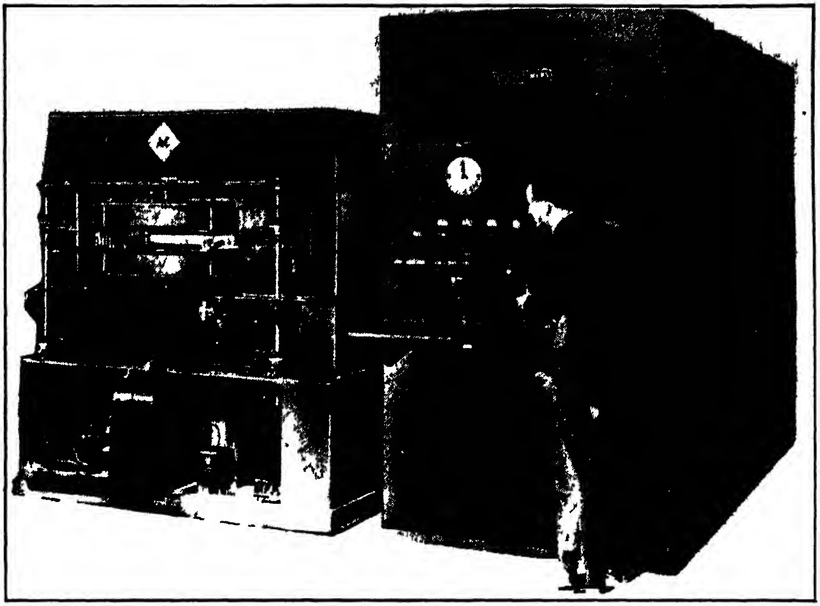


Fig 10 20-mev betatron and control unit (Courtesy Allis-Chalmers Manufacturing Company.)

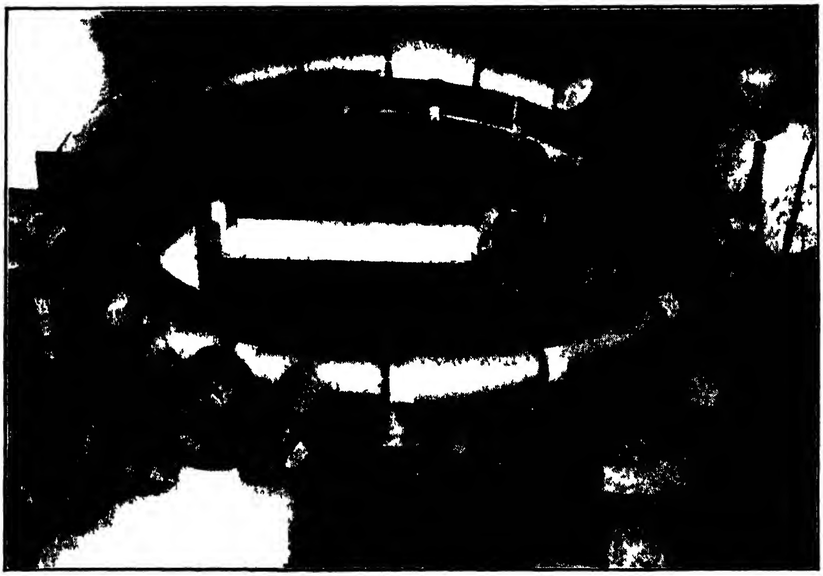


Fig. 11 View of the tube for a 100-mev betatron. (Courtesy General Electric X-Ray Corporation.)

second each electron travels over 800 miles. By means of the pulsing field they may be removed from the circular orbit at any time during the quarter-cycle, delivering energy varying from 1 mev up to the full 100 mev. The X-ray beam output has been as high as 2600 r (roentgens) at 100 mev, dropping to about 7 per cent of this value at 20 mev. A view of the tube for a 100-mev betatron is shown in Fig. 11.

In addition to its application in the radiographic field, the betatron holds much promise for use in both nuclear research and medical therapy. The reader will find detailed explanations of the theory, construction, and operation in the references cited at the end of this chapter.

Frequency Limitations of Electron Tubes. The vacuum tubes considered in the preceding pages of this book are limited to use with low, medium, and high frequencies. (*See* frequency chart on page 430.) The three following limiting factors apply to these tubes.

1. Transit time of electrons.
2. The impedance of the lead inductance and interelectrode capacitance of the tube.
3. The radiation loss of the external circuits at ultra-high frequencies.

The time required for an electron to pass from the cathode to the anode is very short, but at extremely high frequencies this transit time may approach the period for a cycle and then a delay and phase angle is introduced into the amplified signal. The inductance of an external conductor and the capacity between a pair of conductors has little effect upon the resulting voltage and current at medium frequencies, but at high frequencies the resulting reactance presents many design difficulties. It was suggested earlier that radiation of high-frequency carriers is very effective and that at ultra-high frequency the undesired radiation losses from the circuit external to the tube reduce the efficiency of the circuit. These three disadvantages of ordinary electron tubes are overcome in several special tubes such as the lighthouse, the Klystron, and the magnetron.

The Lighthouse and Disk Seal Tube. The lighthouse tube is a high-frequency vacuum tube which utilizes new principles of construction. The elements of its construction and the general assembly of a lighthouse triode are shown in Fig. 12. Part (*a*) shows the anode as the circular base of a metallic cylinder, the grid as a circular fine mesh, and the cathode as an oxide-coated circular top of a hollow cylinder containing a heater. Thus the anode, grid, and cathode are circular sections of planes which may be brought very close together. The

cathode-to-grid spacing is of the order of 4 or 5 mils and the grid-to-plate spacing is 12 or 13 mils. This short spacing reduces the transit time and overcomes the first limitation of earlier tubes. The tube elements are supported by disks which, in turn, are separated, insu-

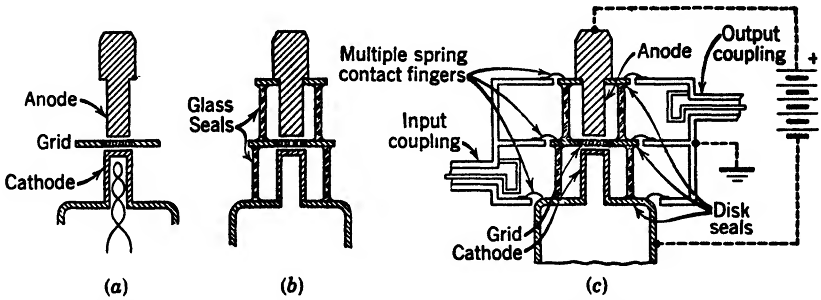


FIG. 12. Parts and assembly of a lighthouse tube and resonators. (Courtesy General Electric Company.)

lated, and sealed to hollow glass cylinders, as shown in part (*b*). The enclosure is evacuated, completing the vacuum tube but not its operating circuit.

The improved features of the operating circuit of the lighthouse tube are illustrated in the steps of Fig. 13. Part (*a*) shows the sche-

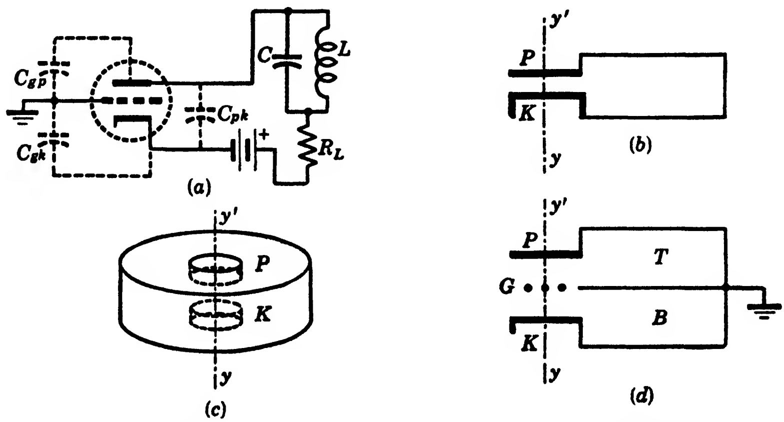


FIG. 13. Steps in theory and development of a cavity resonator.

matic circuit for an ordinary triode using a grounded grid. The impedance of the output or load circuit contains elements of resistance, inductance, and capacitance arising from lumped units plus the inter-electrode and line capacitance and the inductance of the load line itself.

At high frequencies the magnitude of the units of inductance in the load circuit will approach infinity. The resulting impedance can be greatly reduced by changing the cathode plate load to the simple equivalent circuit of part (*b*), Fig. 13. This new impedance can be reduced further in a new structure formed by rotating the line load of part (*b*) about the vertical axis yy'. This concept produces the hollow cylindrical box shown in part (*c*) having a circular plate P and cathode K at top and bottom. Now a vertical electrical field exists within the box, and the plate load electron current flows axially (along yy') from K to P, then radially on the inside surface of the box to the outside, and then returns radially to the cathode. This flow of plate current will induce a magnetic field surrounding the axis yy'. Obviously, the flow of high-frequency plate current on the inside of the chamber finds a reduced impedance over the circuit of Fig. 13*b* and, of equal importance, the circuit has perfect shielding and there is no radiation loss. Thus this structure which is called a cavity or resonance chamber serves to overcome limiting factors number 2 and 3 stated on page 418. This cavity is called a resonance chamber because it constitutes a high-frequency parallel tuned circuit. The circular faces of P and K separated by a vacuum constitute a condenser, and the inside circuit of the chamber constitutes one turn of an inductance having a very low resistance. Obviously, this parallel LC circuit has a high resonant frequency and a very high Q. Resonance is a valuable characteristic of all cavities having this type of construction. The final step in triode construction is suggested in the line circuit of part (*d*) of Fig. 13. If this circuit is rotated about its axis yy', a double cavity resonator is generated in which the cathode-grid input exists in one isolated chamber and the grid-plate output in the upper chamber. The applica-

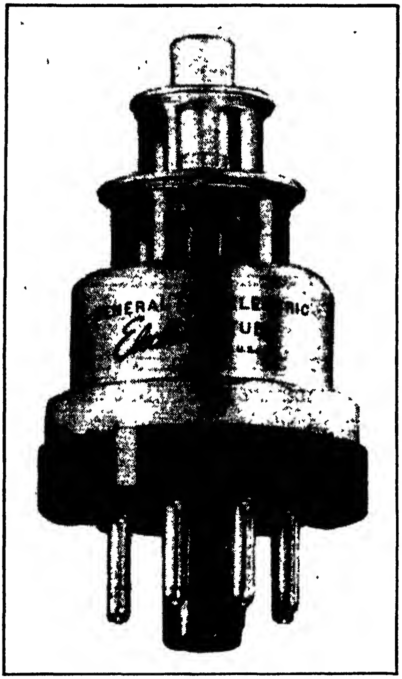


FIG. 14. Lighthouse tube (2C40). (Courtesy General Electric Company.)

tion of these principles of construction is shown in Fig. 12c wherein the tube of part (*b*) has been inserted into a double cavity. In the actual circuit the plate, grid, and cathode disk are insulated by blocking mica disks to prevent a short circuit to direct current. Obviously, these blocking condensers offer little impedance to the circulation of high-frequency currents around the inside of the cavity. The direct current for the cathode plate circuit is connected as indicated. The high-frequency input and output are fed through coaxial cables and are coupled to the circulating magnetic fields by the coil loops as shown for use as an amplifier. For use as an oscillator the input and output cavities are coupled through a coaxial cable or by other arrangements. The input and output cavities may be tuned by mechanical systems which change the volume and length of the resonating current path within the chambers. A commercial form of the lighthouse tube is illustrated in Fig. 14. This tube has an upper frequency limit of about 3500 megacycles.

The Klystron. The Klystron is an ultra-high-frequency oscillating tube of the cavity resonator type which operates on the principle of

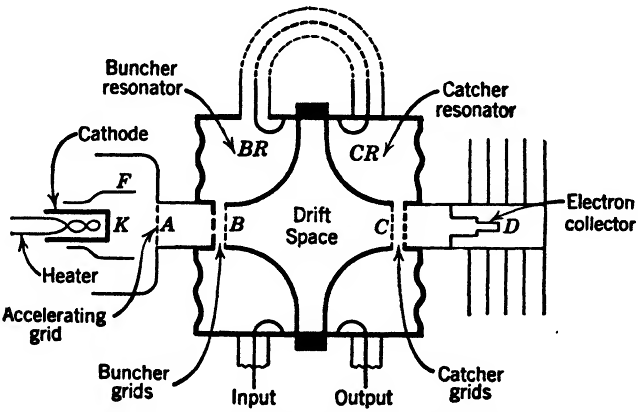


FIG. 15. Cross section of a Klystron.

velocity modulation. The construction of one of these tubes and its principle of operation may be understood by a consideration of the cross-sectional view of Fig. 15. Electrons emitted from the oxide-coated heater cathode K are accelerated by an electron gun consisting of a focusing cap F and an accelerating grid-like anode A. The resulting electron stream moves on the line $ABCD$ through the buncher grids B, catcher grids C, to the electron collector at D. Like the lighthouse tube described in the preceding article, the output of the tube is not measured directly by the electron current flow in the cathode-

anode collector circuit but by high-frequency currents produced in the resonators or cavities. This tube has two resonating (*LC*) cavities. One resonator *BR* is called the buncher resonator and its *LC* circuit appears across the buncher grids *B*. The second or output resonator *CR* is called the catcher resonator and has its *LC* circuit across the catcher grids *C*. Between these two cavity resonators is a third chamber called the drift space. In action any resonating or oscillating current in the buncher resonator *BR* produces a high-frequency potential across the buncher grids *B* which causes the electrons to move into the drift space in "bunches" and continue through the catcher grids in a similar manner. To understand the action taking place it should be noted that the left-hand grid at *B* and the right-hand grid at *C* are held at a constant potential which is ground (Fig. 16). Hence neither of these grids changes potential in action but *their mates do change potential relative to them.* Any high-frequency input into cavity *BR* causes a corresponding high-frequency current in the *BR* cavity which, in turn, varies the potential across the grids at *B*. Thus the potential of the right-hand grid of *B* varies with respect to its mate, and this variation reacts upon the electron stream passing through *B*. When the relative potential is higher (+) the velocity of the passing electrons is increased, and when it is lower (−) the velocity is reduced. The result is to vary the velocity of the electrons entering the drift space. This periodic variation of velocity permits the faster electrons to overtake the slower ones as they move through the drift space so that by the time of arrival at the *C* grids they have formed into bunches of electrons (negative space charge). Now these bunches of moving electrons constitute pulses or "spurts" of current (electrons in motion). As these pulses of current pass through the space between grids *C* they produce circulating pulses of flux in the *CR* cavity. These pulses of flux cause an oscillating current in the *CR* resonant cavity. If these pulses correspond to the resonant frequency of the cavity a relatively large current results from the familiar *LC* action. The passing of the bunches of electrons or pulses of cur-

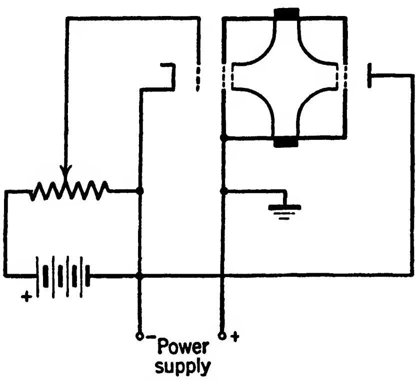


FIG. 16. Direct-current circuit for a Klystron.

rent imparts energy to the resonant cavity by storing energy in its magnetic field. Thus the energy for the *CR* cavity is derived from the electron stream, and the average velocity of the electron is re-

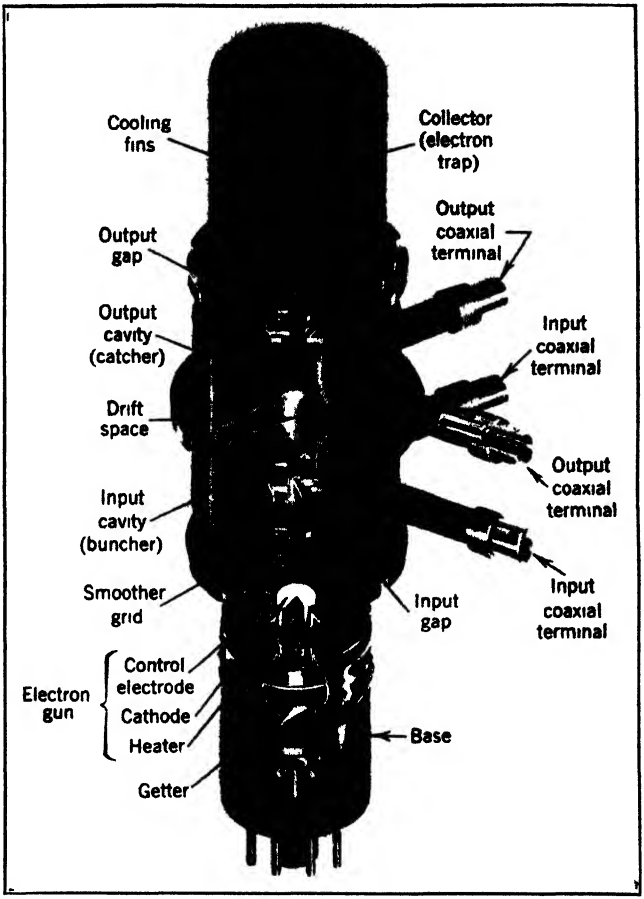


FIG. 17. Commercial double-resonator Klystron (Courtesy Sperry Gyroscope Company)

duced correspondingly. The action of grid *B* which results in bunches in the drift space constitutes a velocity-modulating action. In the course of the action described some electrons land on the grids as well as on the electron collector, and all are returned to the d-c power supply. The schematic circuit for the Klystron is given in Fig. 16. To complete the physical concept of Klystron action it should be noted that the initial source of energy is supplied by the battery or power

supply in the electron gun. A part of the energy given by the electron gun is absorbed by the grids at B, and a much larger amount is absorbed by grids C and their resonating cavity CR. The useful output from the Klystron is tapped from the magnetic field in resonator CR by a loop from a coaxial cable marked output. When the Klystron is used as an oscillator both resonators are tuned to the same frequency and a feedback is provided from the catcher resonator to the buncher resonator by means of a coaxial cable, as shown at the top of Fig. 15. The resonant frequency of either cavity is controlled (tuned) by a variation of the spacing between the grids of B and C. This tuning is accomplished by external mechanical movement permitted by the bellows-like construction of ends of the resonating chamber (see Fig. 15).

A Klystron may be used as a frequency multiplier if the catcher resonator is tuned to a harmonic of the buncher voltage. It may also be used as an amplifier (double resonator type) by feeding the input into the buncher as indicated in Fig. 15. The power outputs obtainable from Klystrons are of the order of a fraction of a watt to several hundred watts. Commercial Klystrons (of various types) operate from 1000 to 24,000 megacycles. A commercial form of double-cavity Klystron is shown in Fig. 17.

The Magnetron. The magnetron is a high-frequency oscillator which utilizes both electric and magnetic fields as its source of operation. An early form of magnetron (1920–1930) is illustrated in Fig. 18. It consists of an axial heater-type cathode surrounded by a split anode and a coaxial coil which produces a magnetic field parallel to the axis and cathode. The circuit for this type (exclusive of magnetic coil and heater) is shown in Fig. 19. With zero current in the magnetic coil the electrons emitted by cathode K will move on radial lines to the split anode. When the magnetic coil is excited the axial magnetic field will deflect the moving electrons and cause them to move in curved paths, as explained on page 16. The curvature of the electron path can

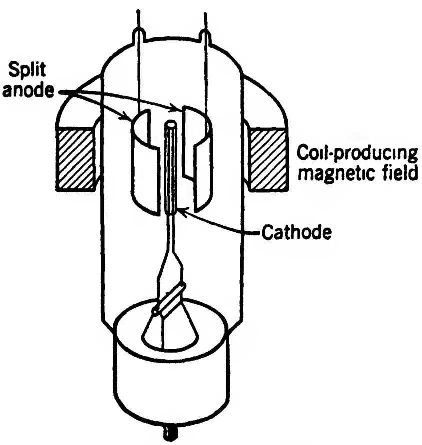


FIG. 18. Split-anode magnetron.

be controlled by adjustment of the magnetic field. A suitable adjustment of the field strength and battery supply voltage will cause the circuit to oscillate at the resonant frequency of the external *LC* tank. The oscillation in the tank circuit varies the voltage across itself and across the split anode *A*1 and *A*2 in such a manner as to maintain oscillation by utilizing power from the d-c supply. As the voltage on anode *A*1 rises it exerts a greater pull on the electrons and changes the curvature of their path. The phase relation of the

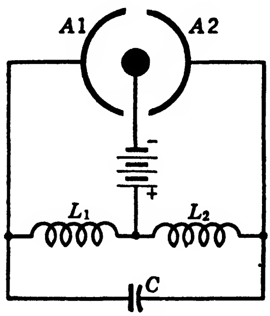


FIG. 19. Oscillator circuit for a split-anode magnetron.

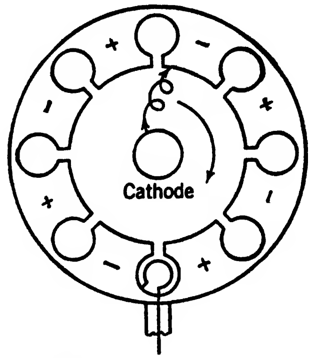


FIG. 20. Cross section of a cavity magnetron.

circuit is such that when the voltage of *A*1 is the maximum its current due to electrons landing on it tends to be the minimum, and vice versa. Thus the magnetron oscillator is of the negative resistance type. This form of magnetron is capable of producing oscillations up to approximately 600 megacycles. Its frequency is limited by the resonant frequency of this form of external tank circuit and by the transit time of the electrons moving from cathode to anode. The output for this oscillator is obtained through a coil coupled to L_1L_2.

Magnetrons for producing much higher frequencies and having much greater outputs were developed for military purposes during World War II. These magnetrons use resonant cavities for the *LC* tank circuit and powerful permanent magnets for producing a magnetic field parallel to the axis of the cathode. A cross section of this magnetron is shown in Fig. 20. The anode consists of a solid copper block machined as illustrated, giving an appearance similar to a multiple field of a d-c dynamo. The circular holes in the anode plus the ducts cut to the inside of the annular structure constitute resonant cavities in

which the high-frequency energy is generated. Electrons emitted from the cathode are attracted to the surrounding anode and are forced to travel in epicycloidal paths by the powerful magnetic field. The changing phase relation of the rf field causes some electrons to return to the cathode while others are moved outward in cycloidal paths

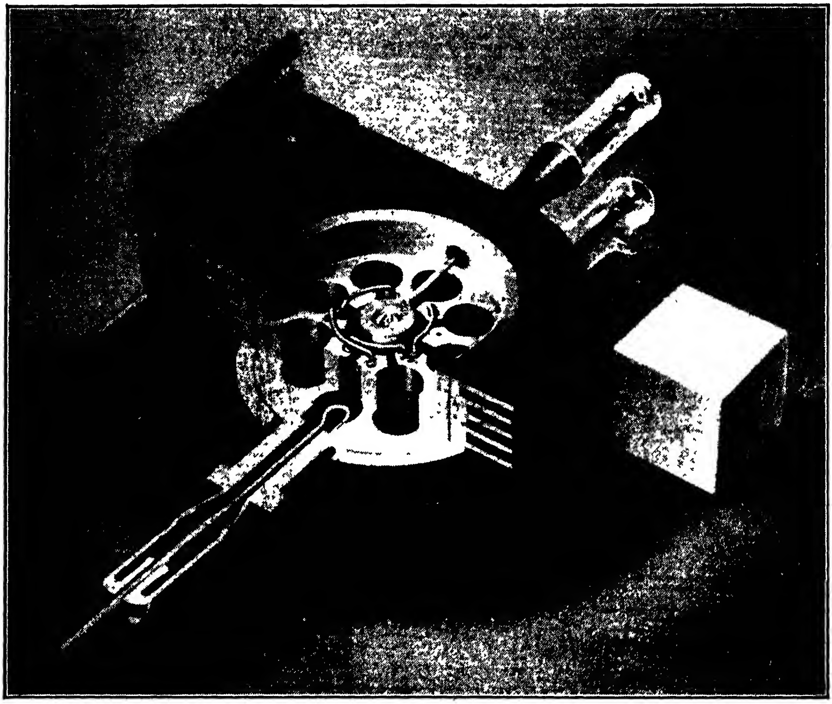


FIG. 21. Sectional view of high-power magnetron. (Courtesy Bell Telephone Laboratories.)

which have a rotary motion within the cathode-anode space. This rotary motion is similar to the rotating magnetic field in the stator of an a-c machine. In passing the gaps in front of the resonator cavities the electronic rotating field imparts energy to these multiple resonators. High-frequency power output is taken from the magnetron by means of a coaxial cable and loop, as indicated at the bottom of Fig. 20. Magnetrons of this type may be "pulsed" by applying periodically a high d-c voltage to the cathode-anode circuit. The frequency range of this device is of the order of 3000 to 30,000 megacycles. A cutaway section of a commercial magnetron is given in Fig. 21.

Problems

1. Explain the operation of the Villard circuit shown in Fig. 4.

2. An X-ray therapy tube operates continuously at 250,000 volts direct current with a current of 10 milliamperes. If 90 per cent of the energy must be conducted away by water cooling, what will be the rate of water flow for a temperature rise (water) of 40 degrees F?

3. If 1,000,000 volts is applied between the cathode and target of an X-ray tube, what would be the velocity of the impinging electron calculated by equation 9 of Chapter I? How does this compare with the velocity of light? Explain.

4. Assume that the "donut" tube in a betatron has a mean diameter of 1 meter and that an electron in the tube gains energy at the average rate of 200 electron-volts for each revolution. What will be its velocity starting from rest after 0.001 second?

REFERENCES

SLEPIAN, J., United States Patent 1,645,304, October 11, 1927.

KERST, D. W., *Phys. Rev.*, **58**, 841 (1940).

KERST, D. W., and R. SERBER, *Phys. Rev.*, **60**, 53 (1941).

GIRARD, J. P., and G. D. ADAMS, *Elec. Eng.*, May 1946.

WESTENDORP, W. F., and E. E. CHARLTON, *J. Applied Phys.*, **16**:10, 581–593 (October 1945).

VACUUM ELECTRONIC TUBES *

Tube Names	Function	Application	Type of Control	Type of Cathode
Kenotron Diode	Rectifies low values of current (often at high voltage)	Rectification in radio, X ray, and electrostatic precipitation	None	Hot
X Ray	Produces X-ray radiation	Medical diagnosis and therapy Examination of industrial products	None	Hot
Pliotron May be triode, tetrode, or pentode	Amplification and high frequency oscillation	High-frequency heating Carrier current Control circuits Radio receivers and transmitters	Electrostatic Grid	Hot
Cathode Ray	Produces visual indications by controlled electron beam	Circuit analysis Television	Electrostatic or Electromagnetic	Hot
Magnetron	Ultra-high-frequency oscillation	Short-wave radio	Electromagnetic field	Hot
Klystron	Ultra-high-frequency oscillation	Used in military radio and detection devices	Resonance chamber	Hot
Vacuum Phototube	Used in applications where audio-frequency response is important	Sound movies	Light variation	Photoelectric

* This table and the one which follows condensed from *Electronics at Work*, Westinghouse Electric Corporation, 1943.

GAS-FILLED ELECTRONIC TUBES

Tube Names	Function	Application	Type of Control	Type of Cathode
Rectigon Tungar	Low voltage rectification	Battery charging	None	Hot
Phanotron Mercury-vapor rectifier	Rectification of moderate amounts of current at voltages up to 20,000	Radio transmitters and industrial applications	None	Hot
Thyratron May be gas triode or gas tetrode	Control and controlled rectification of moderate amounts of current at voltages up to 22,000	Welding control Timing circuits Motor control Voltage regulation	Electrostatic Grid	Hot
Glow Tube	Illumination, rectification, and regulation at low current, low voltage	Switchboard indicating lights Control circuit voltage regulators Protection of supervisory control circuits	None	Cold
Grid-Glow Tube Gas triode	Controls small amounts of power	Safety control of furnaces	Electrostatic Grid	Cold
Pool Tube Mercury-arc rectifier			None	Pool (usually mercury)
Grid-Pool Tube Excitron Mercury-arc rectifiers	Rectification, conversion, and control of large amounts of power	Power for aluminum and magnesium production. Electrochemistry. Transportation. Resistance welding.	Electrostatic Grid	Pool (usually mercury)
Ignitron			Resistance ignition	Pool (usually mercury)
Gas Phototube	Process and quality control Detection devices	Sound movies Pinhole detector Cutting, winding and printing control	Light variation	Photoelectric

FREQUENCY CHART

Designations	Abbreviations	Frequencies	Wavelengths (In Meters)
Radio frequencies	rf		
Very low		10 to 30 kc	30,000 to 10,000
Low		Above 30 to 300 kc	10,000 to 1,000
Medium		Above 300 to 3,000 kc	1,000 to 100
High		Above 3,000 to 30,000 kc	100 to 10
Very high		Above 30,000 to 300,000 kc	10 to 1
Ultra-high		Above 300,000 to 3,000,000 kc	1 to 0.1
Super-high		Above 3,000,000 to 30,000,000 kc	0.1 to 0.01
Audio	af	30 to 15,000 cycles	
Video	vf	30 to 5,000,000 cycles and higher	

INDEX